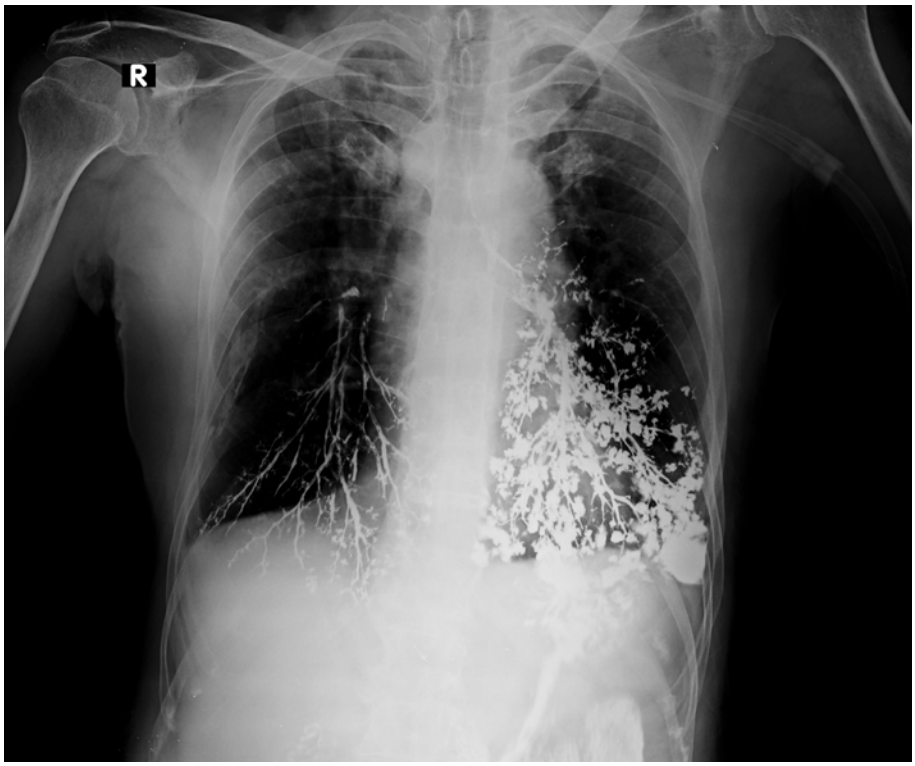
The Netherlands Journal of Medicine

PUBLISHED IN COLLABORATION WITH THE NETHERLANDS ASSOCIATION OF INTERNAL MEDICINE



Breathlessness after a radiological procedure: what is your diagnosis?

May 2016, VOL. 74, NO. 4, ISSN 0300-2977

VAN ZUIDEN COMMUNICATIONS

The Netherlands Journal of Medicine

MISSION STATEMENT

To serve the need of the physician to practice up-to-date medicine and to keep track of important issues in health care. To promote and to enhance clinical knowledge by publishing editorials, original articles, reviews, papers regarding specialty training, medical education and correspondence.

EDITORIAL INFORMATION

Editorial office
Erasmus MC, University Medical Center Rotterdam
Department of Internal Medicine
's-Gravendijkwal 230
3015 CE Rotterdam
The Netherlands
Tel.: +31 (0)10-703 59 54
Fax: +31 (0)10-703 32 68
E-mail: p.l.a.vandaele@erasmusmc.nl
http://mc.manuscriptcentral.com/nethjmed

CITED IN

Biosis database; embase/excerpta medica; index medicus (medline) science citation index, science citation index expanded, isi alerting services, medical documentation services, current contents/clinical medicine, PubMed.

ISSN: 0300-2977

VAN ZUIDEN communications

Van Zuiden Communications B.V.
PO Box 2122
2400 CC Alphen aan den Rijn
The Netherlands
Tel.: +31 (0)172-47 61 91
Email: mouton@vanzuidencommunications.nl
Internet: www.njm-online.nl

Contents

EDITORIAL

Total or free, that is the question

F.H. de Jong

Department of Internal Medicine, Endocrine Laboratory, Erasmus university Medical Centre, Rotterdam, the Netherlands, email: f.h.dejong@erasmusmc.nl

In this issue of the Netherlands Journal of Medicine, Vastbinder et al.[1] describe the effect of oestrogen-containing oral contraceptives on the results of the dexamethasone screening test (DST) in normal women. The DST is one of the tests used to exclude the diagnosis of Cushing's syndrome. The test erroneously showed cortisol levels above the cut-off value of 50 nmol/l in 8 out of 13 women, presumably due to increased levels of the liver-produced glycoprotein cortisol-binding globulin (CBG), which are stimulated by the oestrogenic component of the contraceptive pills. Results of the DST normalised one week after oestrogen withdrawal in all but one of the women, and in all women after another five weeks.

The authors did not estimate actual levels of CBG, presumably because the mechanism leading to the increased CBG levels and therefore to the disturbed DST result after oestrogen administration is well known.[2] However, a number of other factors, as mentioned below, may also affect the results of measurements of protein-bound hormones and the interpretation of related function tests. Many of these depend on changes in the level of binding proteins. This has clearly been recognised to be the case for thyroid hormones; measurement of free thyroxin has superseded the estimation of total thyroxin for a long time already. The situation for steroid hormones is different, probably because estimation of free steroid concentrations by dialysis is time consuming, direct assays for serum free steroids are notoriously unreliable, and the number of available CBG assays is limited. Nevertheless, there are possibilities to overcome these problems, which will also be discussed below.

FACTORS AFFECTING CONCENTRATIONS AND EFFECTS OF CBG

Circulating cortisol is partly bound to CBG (ca 70%), to albumin (ca 10-15%) and is partly free in the circulation. Biological effects are presumably exerted by the sum of albumin-bound and free cortisol (i.e. non-CBG-bound cortisol), since the binding to albumin has a low affinity and is readily disrupted while the CBG-cortisol complex will not be separated during the time necessary to pass the vascular bed of a target organ. In dialysis experiments, only free cortisol is measured.

Apart from the above-mentioned stimulating effect of oestrogens on serum CBG levels, CBG concentrations can be suppressed by increased levels of immune modulators such as interleukin-6, by insulin, thyroxin, by growth hormone treatment through increased levels of IGF-1 and by liver disease, e.g. cirrhosis.[3] All of these conditions may lead to erroneously low levels of total cortisol suggesting hypocortisolaemia or falsely suppressed cortisol in DSTs. Furthermore, mitotane treatment for adrenal cortical carcinoma will stimulate CBG levels, possibly leading to misinterpretation of total cortisol levels in patients with this disease.

A different reason for misinterpretation of total serum cortisol concentrations may be the presence of mutations in the gene coding for this protein[3] leading to suppressed total, but normal non-CBG-bound or free cortisol levels in serum. Interestingly, one of these mutations is prevalent in Han Chinese, where it leads to a so far unexplained preference for female offspring.[4] Finally, increased levels of total and free cortisol in the absence of signs and symptoms of hypercortisolaemia can be found in patients with mutations in the gene coding for the glucocorticoid receptor.[5]

ALTERNATIVES FOR DIRECT ESTIMATION OF FREE CORTISOL

Apart from the measurement of serum levels of free cortisol by dialysis, an approximation can be made by calculation, using the levels of total cortisol, CBG and albumin in the formula described by Dorin et al.[6] Alternatively, salivary levels of cortisol are strongly correlated with free cortisol levels in serum samples taken at the same time,[7] whereas the cortisol level measured in hair reflects the mean serum free cortisol concentration as present during a longer period.[8] Assuming a hair

growth rate of 1 cm/month, it is possible to study changes in mean cortisol levels during illness or periods of stress by measuring cortisol in subsequent centimetres of hair.

TOTAL TESTOSTERONE CONCENTRATIONS ARE STRONGLY AFFECTED BY SHBG LEVELS

Like CBG, sex hormone-binding globulin (SHBG) is a glycoprotein, produced and secreted by the liver, which binds testosterone, 5α-dihydrotestosterone and oestradiol. Approximately 50% of testosterone is SHBG-bound in men; in women the SHBG-binding amounts to 80%. Total testosterone levels in men are strongly related with single nucleotide polymorphisms (SNPs) in the SHBG gene[9] and with the serum level of SHBG,[10] whereas the concentration of serum non-SHBG-bound testosterone is independent of the SHBG level. In its turn, the SHBG concentration is partly dependent on SNPs in a larger number of other genes, which encompass multiple biological pathways, including hepatic function, lipid metabolism, carbohydrate metabolism, type 2 diabetes, and androgen and oestrogen receptor function.[11] These observations are in line with earlier findings on direct effects of oestrogens, androgens and insulin on SHBG levels, and follow a similar pattern compared with the factors affecting CBG concentrations. Finally, one case of an inactivating mutation in the SHBG gene has been described in a man with an inadequately low level of total testosterone but normal gonadal development and spermatogenesis.[12]

ALTERNATIVES FOR DIRECT ESTIMATION OF FREE TESTOSTERONE

Similar to the situation for cortisol, methodologies for the calculation of free or non-SHBG bound testosterone have been developed. However, the concordance between the various methods was only limited,[13] indicating that valid conclusions can only be drawn from comparisons with reference values obtained using the same method. A much simpler approximation of the concentration of non-SHBG-bound testosterone is the calculation of the free androgen index, defined as total testosterone x100/SHBG, where concentrations of both testosterone and SHBG are expressed as nmol/l. This method might yield meaningful results in women, where total testosterone levels are much lower than SHBG concentrations.[14] However, in men, where testosterone levels exceed SHBG concentrations by a factor between 1.5 and 2, this will not lead to meaningful results. Relatively new developments are estimations of testosterone in saliva[15] and hair,[16] which might also reflect the serum concentration of non-SHBG-bound testosterone.

CONCLUSIONS

Changes in the concentrations of specific steroid-binding proteins or in the affinity of their binding to steroids will become visible in the total concentration of the steroid, while the non-protein bound concentration will only be affected slightly. For this reason, if unexpected results of determinations of steroid hormones are encountered, it may be possible to resolve these discrepancies by investigation of the quantity and quality of the specific steroid-binding proteins.

REFERENCES

1. Vastbinder M, Kuindersma M, Mulder AH, Schuijt MP, Mudde AH. The influence of oral contraceptives on overnight 1 mg dexamethesone suppression test. Neth J Med. 2016;74:XXXX.
2. Nickelsen T, Lissner W, Schöffling K. The dexamethasone suppression test and long-term contraceptive treatment: measurement of ACTH or salivary cortisol does not improve the reliability of the test. Exp Clin Endocrinol. 1989;94:275-80.
3. Gagliardi L, Ho JT, Torpy DJ. Corticosteroid-binding globulin: the clinical significance of altered levels and heritable mutations. Mol Cell Endocrinol. 2010;316:24-34.
4. Lei JH, Yang X, Peng S, et al. Impact of corticosteroid-binding globulin deficiency on pregnancy and neonatal sex. J Clin Endocrinol Metab. 2015;100:1819-27.
5. Lamberts SW, Huizenga AT, de Lange P, de Jong FH, Koper JW. Clinical aspects of glucocorticoid sensitivity. Steroids. 1996;61:157-60.
6. Dorin RI, Pai HK, Ho JT, et al. Validation of a simple method of estimating plasma free cortisol: role of cortisol binding to albumin. Clin Biochem. 2009;42:64-71.
7. Dorn LD, Lucke JF, Loucks TL, Berga SL. Salivary cortisol reflects serum cortisol: analysis of circadian profiles. Ann Clin Biochem. 2007;44:281-4.
8. Manenschijn L, Koper JW, van den Akker EL, et al. A novel tool in the diagnosis and follow-up of (cyclic) Cushing's syndrome: measurement of long-term cortisol in scalp hair. J Clin Endocrinol Metab. 2012;97:E1836-43.
9. Ohlsson C, Wallaschofski H, Lunetta KL, et al. Genetic determinants of serum testosterone concentrations in men. PLoS Genet. 2011;7:e1002313.
10. De Ronde W, van der Schouw YT, Muller M, et al. Associations of sex-hormone-binding globulin (SHBG) with non-SHBG-bound levels of testosterone and estradiol in independently living men. J Clin Endocrinol Metab. 2005;90:157-62.
11. Coviello AD, Haring R, Wellons M, et al. A genome-wide association meta-analysis of circulating sex hormone-binding globulin reveals multiple Loci implicated in sex steroid hormone regulation. PLoS Genet. 2012;8:e1002805.
12. Vos MJ, Mijnhout GS, Rondeel JM, Baron W, Groeneveld PH. Sex hormone binding globulin deficiency due to a homozygous missense mutation. J Clin Endocrinol Metab. 2014;99:E1798-802.
13. De Ronde W, van der Schouw YT, Pols HA, et al. Calculation of bioavailable and free testosterone in men: a comparison of 5 published algorithms. Clin Chem. 2006;52:1777-84.
14. Daan NM, Jaspers L, Koster MP, et al. Androgen levels in women with various forms of ovarian dysfunction: associations with cardiometabolic features. Hum Reprod. 2015;30:2376-86.
15. Büttler RM, Peper JS, Crone EA, Lentjes EG, Blankenstein MA, Heijboer AC. Reference values for salivary testosterone in adolescent boys and girls determined using Isotope-Dilution Liquid-Chromatography Tandem Mass Spectrometry (ID-LC-MS/MS). Clin Chim Acta. 2016;456:15-8.
16. Noppe G, de Rijke YB, Dorst K, van den Akker EL, van Rossum EF. LC-MS/MS-based method for long-term steroid profiling in human scalp hair. Clin Endocrinol (Oxf). 2015;83:162-6.

REVIEW

T-cells fighting B-cell lymphoproliferative malignancies: the emerging field of CD19 CAR T-cell therapy

D.M. Heijink[1], A.P. Kater[1], M.D. Hazenberg[1], A. Hagenbeek[1], M.J. Kersten[1]*

[1]Department of Hematology, Academic Medical Center, Amsterdam, the Netherlands, *corresponding author: tel.: +31 (0)20 - 56 65 785, fax: +31 (0)20 - 69 19 743, email: m.j.kersten@amc.nl

ABSTRACT

CAR T-cells are autologous T-cells transduced with a chimeric antigen receptor (CAR). The CAR contains an antigen recognition part (originating from an antibody), a T-cell receptor transmembrane and cytoplasmic signalling part, and one or more co-stimulatory domains. While CAR T-cells can be directed against any tumour target, most experience thus far has been obtained with targeting of the B-cell antigen CD19 that is expressed by B-cell acute lymphocytic leukaemia, chronic lymphocytic leukaemia and other B-cell lymphomas. The first clinical results are promising, although there are profound differences in response between patients with different haematological malignancies. Treatment-related side effects have been observed that require specific management. This review will explain the mechanism of action, summarise the experience to date and point out future directions for this hopeful new addition to the therapeutic armamentarium in the treatment of lymphoproliferative B-cell malignancies.

KEYWORDS

CAR T-cells, CD19, B-cell malignancies

INTRODUCTION

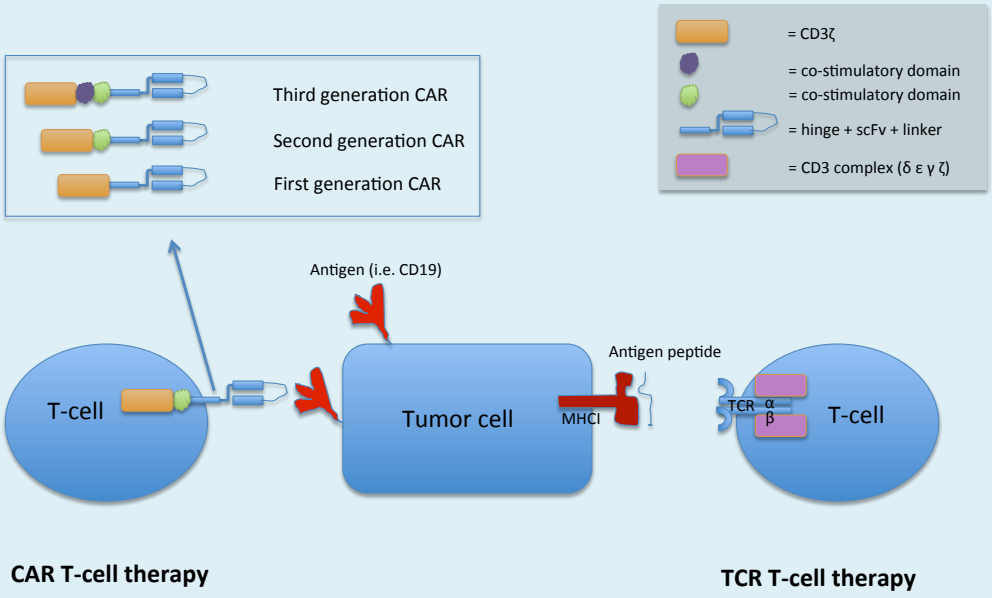
The cornerstone of the treatment of lymphoproliferative malignancies consists of chemotherapy, regularly followed by autologous or allogeneic stem cell transplantation. These treatments are not tumour cell specific and cause many side effects. In the last decade the focus has moved toward so-called 'targeted therapies' using monoclonal antibodies and/or small molecules. Important examples in the current therapeutic arsenal are monoclonal antibodies against CD20, CD52 and CD30, cell surface antigens frequently, but not exclusively, expressed by lymphoproliferative malignancies such as lymphoma and lymphatic leukaemia. Other examples include small molecules targeting intracellular signalling pathways important for growth and survival of malignant cells, such as JAK inhibitors, BTK inhibitors, PI3K inhibitors and inhibitors of the proteasome. Most recently, an effective inhibitor of BCL-2, which induces apoptosis of malignant cells, has become available.[1]

Another area that is attracting a lot of attention concerns autologous anti-tumour immune responses, in particular by T-lymphocytes. Most tumours are immunogenic but anti-tumour T-cell responses are efficiently suppressed by the tumour through membrane expression of immunosuppressive proteins such as PD-1 ligand and CTLA4 ligand. Interaction of these proteins with PD-1 and CTLA-4 on tumour specific T-cells renders these T-cells anergic. Antibodies that target PD-1 and CTLA-4 interactions (e.g. nivolumab, ipilimumab), release the breaks on the anti-tumour immune response and have been proven very effective.[2] Another example is blinatumomab, a bispecific T-cell engager (BiTe). This fusion protein consists of a CD3-specific part (recognising T-cells) and a CD19 specific part (recognising B-cells) and can thereby redirect T-cells directly towards malignant B-cells. Blinatumomab is approved for refractory or relapsed B-cell acute lymphocytic leukaemia (B-ALL).[3] Even more sophisticated is the use of engineered CAR T-cells, the topic of this review.

ENGINEERED T-CELL THERAPY IN HAEMATOLOGY

This completely new field involves the administration of autologous T-cells that have been engineered ex vivo to recognise tumour cells. Roughly, there are two ways

Figure 1. *Schematic representation of CAR T-cell and TCR T-cell therapy*



to employ T-cells against haematological malignancies: T-cell receptor (TCR) engineered T-cells or chimeric antigen receptor (CAR) T-cells *(figure 1)*. TCR T-cells are autologous T-cells transferred with a gene for the α and β chain of a specific T-cell receptor in order to direct the cells towards tumour-associated antigens (for example gp100 and MART-1), cancer germline antigens (for example NY-ESO-1 and MAGE) or cancer-specific mutations (for example in erbb2). TCR T-cells require a specific human leukocyte antigen (HLA) molecule for recognition of the target antigen. Therefore not all individuals are eligible for this treatment. Today, no results from clinical trials using TCR T-cells in haematological malignancies are known; however, several studies targeting the leukaemia antigen Wilms Tumor-1 (WT-1) are on-going. The great advantage of CAR T-cells is that tumour cell recognition can occur without MHC/HLA restriction. Evidence of clinical efficacy in lymphoproliferative malignancies is piling up quickly.[4-6]

CAR T-CELLS

CAR T-cells are characterised by four important elements *(figure 1)*:

1. A single chain variable fragment (scFv) derived from a murine or humanised antibody which functions as the targeting domain directed against the target of choice. The scFv is attached to an extracellular spacer domain to ensure optimal binding freedom;
2. A hinge and transmembrane domain (often derived from CD8 α or CD28);
3. An intracellular signalling domain that is derived from the TCR (CD3 ζ chain) ensuring T-cell activation;
4. One or two co-stimulatory domains derived from CD28 and/or 4-1BB/CD137. These co-stimulatory domains have been included in the second- and third-generation CAR T-cells, respectively, in order to improve activation, replication and persistence of the CAR T-cells in vivo.

Most of the research to date has focused on autologous CAR T-cells targeting CD19. CD19 is a transmembrane glycoprotein which is expressed on cells of the B-cell lineage, from early pro-B-cells in the bone marrow to mature B-cells in blood and tissues, but not on haematopoietic stem cells. Since CD19 is also expressed on the majority of malignant B-cells, the therapeutic potential of CD19-CAR T-cells is broad.[7,8]

GENERATION OF CAR T-CELLS

The first step in the generation of autologous CAR T-cells is the collection of peripheral blood mononuclear cells through leukapheresis. The cells are cultured in the presence of a T-cell specific mitogenic stimulus, often using beads coated with anti-CD3/anti-CD28 monoclonal antibodies, resulting in the activation of T-cells with high replicative capacity.[7,8] The second step is transduction or transfection of the CAR construct, which is directed

against the target antigen of choice. Multiple methods have been tried including transposon-based systems introducing plasmid DNA into T-cells or electroporation of mRNA into T-cells, but retroviral or lentiviral transduction is currently the transduction mode of choice. CAR T-cells are then cultured in the presence of cytokines (i.e. IL-2, IL-7, IL-12, IL-15). After expansion for several days, the cells are re-infused into the patient. In almost all CD19 CAR T-cell studies patients are lymphodepleted before infusing the engineered T-cells. In B-ALL patients the conditioning therapy mostly consists of fludarabine and cyclophosphamide, in lymphoma patients the choice depends on previous individual treatments. Pre-infusion lymphodepletion facilitates expansion and persistence of CAR T-cells in vivo.[7-10]

CLINICAL USE OF CD19 CAR T-CELLS

After preclinical evaluation, CD19 CAR T-cells were first tested in children and adults with relapsed/refractory ALL in adults with advanced relapsed and refractory chronic lymphatic leukaemia (CLL). Later patients with B-cell non-Hodgkin lymphomas were included as well. *Table 1* demonstrates the phase 1-2 trials which have been performed and published so far, each including ten or more patients.[11-23] Several smaller trials have been performed as well, which are reviewed in other papers.[24,25] Taken together, clear and deep responses are seen in a substantial percentage of B-ALL patients (complete responses in 70-94% of the cases). Durable complete responses correlate with persistence of the CAR T-cells and lasting B-cell depletion. It remains to be determined to what extent CD19 CAR-T cells may be curative for B-ALL and whether it has the potential to replace allogeneic stem cell transplantation in the future. More detailed studies will shed light on the causes of the disappearance and exhaustion of CAR T-cells that heralds B-ALL relapse.[13-15,18,21,23-25]

Data on patients with diffuse large cell B-cell lymphomas are still scarce but so far complete response rates range from 20-100%.[19,20,22,25]

Strikingly, patients with CLL appear to have lower response rates in the generally small trials performed (complete responses in 20-75% of the patients).[16,17,19,22,25] Although not fully elucidated yet, this might be due to disease-specific properties of CLL, in particular the interaction of CLL cells with the tumour microenvironment, which may cause exhaustion and dysfunction of (CAR) T-cells.[26]

In other types of B-cell lymphomas (mantle cell lymphoma, follicular lymphoma) still only limited data are available.

THE DOWNSIDES OF CD19 CAR T-CELLS

There are several serious side effects of CAR T-cell therapy, which are mostly on-target toxicities. CD19 CAR T-cells specifically target CD19 positive B-cells, including healthy B-cells, resulting in long lasting B-cell aplasia. This aplasia – a surrogate for the persistence of CAR T-cells – puts successfully treated patients at risk for infections. Intravenous immunoglobulin therapy can prevent the majority of these infections. In addition, many patients develop cytokine release syndrome with symptoms ranging from mild and flu-like symptoms to severe reactions and multi-organ failure. The latter occurs in 30% of patients, typically 5-21 days after the infusion of CAR T-cells and is mediated by elevated levels of circulating cytokines, especially IL-6, IL-10 and interferon gamma. Treatment with the IL-6 inhibitor tocilizumab is highly effective; however, it is unclear whether this treatment affects the anti-tumour response. The same holds true for treatment with steroids. Finally, a number of patients have developed neurological symptoms ranging from seizures to aphasia and delirium. This type of toxicity seems transient and its pathophysiology is ill-understood.[7,8]

FUTURE DIRECTIONS OF CAR T-CELL THERAPY

Current research focuses on improving the safety and efficacy of CAR T-cell therapy. To be able to switch off the CAR T-cell response, for example in case of excessive toxicity, 'suicide genes' have been incorporated into the T-cells. Examples are introduction of thymidine kinase, which can be inhibited by ganciclovir, expression of inducible caspase or Fas (pro-apoptotic genes) or expression of surface proteins rendering CAR T-cells susceptible to existing agents such as rituximab or cetuximab. Making CAR expression inducible instead of permanent could also increase safety.[4,27]

To increase efficacy, multiple approaches are being explored. Firstly, it is investigated whether specific T-cell subsets provide superior antitumour efficacy. Currently, αβ T-cells (expressing a TCR consisting of an α and β chain) are used; however, therapy with γδ T-cells, specifically the Vγ9Vδ2 subset, may be more effective as these cells specifically recognise and efficiently kill cells with increased levels of phosphoantigens, which have been found in the majority of B-cell lymphomas.[28] Furthermore, it has become clear that selecting less differentiated T-cells such as naïve T-cells, T-memory stem cells and central memory T-cells for transduction results in more effective

Table 1. *Overview of clinical trials using CD19 CAR T-cells in lymphoproliferative malignancies*

Author	Year	ALL (n)	CLL (n)	DLBCL (n)	Other (n)	Co-stimulation	Construct	Clinical response
Maude et al.[11]	2014	30				4-1BB	Lentiviral	27/30 (90%) CR
Grupp et al.[12]	2015	53				4-1BB	Lentiviral	50/53 (94%) CR (NB Extension of study reported by Maude) [13]
Davila et al.[13]	2014	16				CD28	Retroviral	14/16 (88%) CR
Porter et al.[14]	2014		23			4-1BB	Lentiviral	5/23 (22%) CR; 4/23 (17%) PR
Porter et al.[15]	2015		14			4-1BB	Lentiviral	4/14 (29%) CR; 4/14 (29%) PR
Lee et al.[16]	2015	20		1		CD28	Retroviral	ALL: 14/20 (70%) CR
Kochenderfer et al.[17]	2015		4	7	2 indolent	CD28	Retroviral	CLL: 3/4 (75%) CR; 1/4 (25%) PR DLBCL: 4/7 (57%) CR; 2/7 (29%) PR IL: 1/2 (50%) CR; 1/2 (50%) PR
Schuster et al.[18]	2015			13	7 FL, 2 MCL	4-1BB	Lentiviral	DLBCL: 7/13 (54%) ORR FL: 7/7 (100%) ORR MCL: 1/2 (50%) ORR
Turtle et al.[19]	2015	18				4-1BB	Lentiviral	15/18 (= 83%) CR
Turtle et al.[20]	2015		6	18	6 FL, 4 MCL	4-1BB	Lentiviral	CLL: 3/6 (50%) CR; 1/6 (17%) PR DLBCL/FL/MCL: ORR 50 and 67% without and with fludarabine addition
Park et al.[21]	2015	43				CD28	Retroviral	36/43 (84%) CR
Kebriaei et al.[22]	2015	18			3 NHL	CD28	Transposon	ALL plus CLL: 10/21 (48%) CR (NB donor derived T-cells)
Brudno et al.[23]	2015	5	5	5	5 MCL	CD28	Retroviral	ALL: 4/5 (80%) CR CLL: 1/5 (20%) CR, 1/5 (20%) PR DLBCL: 1/5 (20%) CR MCL: 1/5 (20%) PR (NB donor derived T-cells)

CLL = chronic lymphocytic leukaemia; ALL = acute lymphoblastic leukaemia; DLBCL = diffuse large B cell lymphoma; FL = follicular lymphoma; MCL = mantle cell lymphoma; NHL = non-Hodgkin lymphoma; PR = partial response; CR = complete response; ORR = overall response rate; IL = indolent lymphoma.

in vivo expansion and survival.[4] To render CAR T-cells less susceptible to the suppressing effects of regulatory T-cells, T-cell signalling has been altered by increasing the expression of AKT or by modifying the CD28 signalling domain.[28] Interestingly, such an approach could also be of value in improving CAR T-cell therapy in CLL, as T-cells are globally dysfunctional in CLL with a few notable exceptions that could be enforced.[29]

A second approach to improve efficacy is to develop CAR T-cells that express transgenes for IL-2, IL-12, IL-15 or IL-21. These cytokines are important for immune stimulation and may 'armour' the infused CAR T-cells. In addition, chemokine receptors such as IL-7Rα have been built in to improve CAR T-cell survival.

Thirdly, dual or polyspecific CAR T-cells might create higher tumour specificity. An example is the development of CAR T-cells targeting both CD19 and CD5 for CLL.[27] Furthermore, combination therapy with, for example, immune checkpoint inhibitors (inhibiting PD-1 and CTLA4) is being piloted which has the potential to overcome inhibition of the CAR T-cell response.[27,30] To make CAR T-cells more readily available ('off the shelf') the use of allogeneic CAR T-cells is being explored. This obviously poses additional risks of graft-versus-host-disease, which can be overcome by knockdown of the autologous T-cell receptor or introduction of a suicide gene.[29]

In the near future, several clinical trials with CAR T-cells will also become available in the Netherlands (for more information see www.hovon.nl/werkgroepen/llpc).

CONCLUSION

CAR T-cells provide a promising new type of treatment for lymphoproliferative malignancies. Clinical trials in patients with relapsed or refractory B-ALL and different types of B-cell lymphomas demonstrate impressive and often durable responses. The current clinical experience in haematology thus far mostly involves the use of CAR T-cells targeting the B-cell antigen CD19. Developments in novel and more specific target identification, dual target approaches, T-cell subset selection and CAR T-cell engineering will make this treatment even more safe, specific and effective. Hopefully, these developments will bring adoptive T-cell therapies further forward with the potential to replace stem cell transplantation in patients with lymphoproliferative malignancies.

DISCLOSURES

M.J. Kersten and A. Hagenbeek have received financial compensation for advisory boards for Novartis.

REFERENCES

1. https://www.centerwatch.com/drug-information/fda-approved-drugs/therapeutic-area/6/hematology. List of FDA approved drugs for hematology.
2. Mahoney KM, Rennert PD, Freeman GJ. Combination cancer immunotherapy and new immunomodulatory targets. Nature Rev. 2015;14:561-84.
3. Rogala R, Freyer CW, Ontiveros EP, et al. Blinatumomab: enlisting serial killer T-cells in the war against haematological malignancies. Expert Opin Biol Ther. 2015;15:895-908.
4. Bonini C, Mondino A. Adoptive T-cell therapy for cancer: the era of engineered T cells. Eur J Immunol. 2015;45:2457-69.
5. Fujiwara H. Adoptive immunotherapy for haematological malignancies using T cells gene-modified to express tumor antigen-specific receptors. Pharmaceuticas. 2014;7:1049-68.
6. Uttenthal BJ, Chua I, Morris EC, Strauss HJ. Challenges in T cell receptor gene therapy. J Gene Med. 2012;14:386-99.
7. Ghorashian S, Pule M, Amrolia P. CD19 chimeric antigen receptor T cell therapy for haematological malignancies. Br J Haematol. 2015;169:463-78.
8. Maude S, Barrett DM. Current status of chimeric antigen receptor therapy for haematological malignancies. Br J Haematol. 2016;172:11-22.
9. Mato A, Porter DL. A drive through cellular therapy for CLL in 2015: allogeneic cell transplantation and CARs. Blood. 2015;126:478-85.
10. Kim JV, Latouche J-B, Rivière I, Sadelain M. The ABC of artificial antigen presentation. Nature Biotech. 2004;4:403-10.
11. Maude SL, Frey N, Shaw PA, et al. Chimeric antigen receptor T-cells for sustained remissions in leukaemia. N Engl J Med. 2014;371:1507-17.
12. Grupp SA, Maude SL, Shaw PA, et al. Durable remissions in children with relapsed/refractory treated with T-cells engineered with a CD19 targeted chimeric antigen receptor (CTL019). Abstract 681, ASH 2015.
13. Davila ML, Rivière I, Wang X, et al. Efficacy and toxicity management of 19-28z CAR T cell therapy in B cell acute lymphoblastic leukaemia. Sci Transl Med. 2014;19;6:224ra25.
14. Porter DL, Freq NV, Melenhorst JJ, et al. Randomized, phase 2 dose optimization study of chimeric antigen receptor modified T cells directed against CD19 (CTL019) in patients with relapsed, refractory CLL. Blood 2014, 124 (abstract).
15. Porter DL, Hwang WT, Frey NV, et al. Chimeric antigen receptor T cells persist and induce sustained remissions in relapsed refractory chronic lymphocytic leukaemia. Sci Trans Med. 2015;7:303.
16. Lee DW, Kochenderfer JN, Stetler-Stevenson M, et al. T cells expressing CD19 chimeric antigen receptors for acute lymphoblastic leukaemia in children and young adults: 1 phase 1 dose-escalation trial. Lancet. 2015:385:517-28.
17. Kochenderfer JN, Dudley ME, Kassim SH, et al. Chemotherapy-refractory diffuse large B-cell lymphoma and indolent B-cell malignancies can be effectively treated with autologous T cells expressing an anti-CD19 chimeric antigen receptor. J Clin Oncol. 2015;33:540-9.
18. Schuster SJ, Svoboda H, Nasta S, et al. Phase II1 trial of chimeric antigen receptor modified T cells directed against CD19 (CTL019) in patients with relapsed or refractory CD19+ lymphomas. J Clin Oncol. 2015;33 (suppl; abstract 8516).
19. Turtle CJ, Berger C, Sommermeyer D, et al. Immunotherapy with CD19-specific chimeric antigen receptor (CAR)-modified T cells of defined subset composition. J Clin Oncol. 2015;33 (suppl; abstract 3006).
20. Turtle CJ, Berger C, Sommermeyer D, et al. Anti-CD19 chimeric antigen receptor modified T cell therapy for B cell non-Hodgkin lymphoma and chronic lymphocytic leukaemia: fludarabine and cyclophosphamide lymphodepletion improves in vivo expansion and persistence of CAR-T cells and clinical outcomes. Abstract, ASH 2015
21. Park JH, Rivière I, Wang X, et al. Implications of minimal residual disease negative complete remission (MRD-CR) and allogeneic stem cell transplant on safety and clinical outcome of CD19-targeted 19-28z CAR modified T cells in adult patients with relapsed, refractory B-cell ALL. Abstract 682, ASH 2015.
22. Kebriaei P, Ciurea SO, Huls MH, et al. Pre-emptive donor lymphocyte infusion with CD19-directed, CAR-modified T cells infused after allogeneic hematopoietic cell transplantation for patients with advanced CD19+ malignancies. Abstract 862, ASH 2015.
23. Brudno JN, Somerville R, Shi V, et al. Allogeneic T-cells expressing an anti-CD19 chimeric antigen receptor cause remissions of B-cell malignancies after allogenetic hematopoietic stem cell transplantation without causing graft-versus-host disease. Abstract 99, ASH 2015.
24. Maude SL, Teachey DT, Porter DL, Grupp SA. CD-19 targeted chimeric antigen receptor T-cell therapy for acute lymphoblastic leukaemia. Blood. 2015;125:4017-23.
25. Zhu Y, Tan Y, Ou R, et al. Anti-CD19 chimeric antigen receptor-modified T-cells for B-cell malignancies: a systemic review of efficacy and safety in clinical trials. Eur J Haematol. 2015 Jun 25. doi: 10.1111/ejh.12602. [Epub ahead of print]
26. Kater AP, Tonino SH, Egle A, Ramsay AG. How does lenalidomide target the chronic lymphocytic leukaemia microenvironment? Blood. 2014;124:2184-9.
27. Pegram HJ, Park JH, Brentjens RJ. CD28z CARs and armored CARs. Cancer J. 2014;20:127-33.
28. Deniger DC, Switzer K, Mi T, et al. Bispecific T-cells expressing polyclonal repertoire of endogenous γδ T-cell receptors and introduced CD19-specific chimeric antigen receptor. Mol Ther. 2013;21:638-47.
29. Te Raa GD, Pascutti MF, Garcia-Vallejo JJ, et al. CMV-specific CD8+ T-cell function is not impaired in chronic lymphocytic leukaemia. Blood. 2014;123:717-24.
30. June CH, Riddell SR, Schumacher TN. Adoptive cellular therapy: a race to the finish line. Sci Transl Med. 2015;7:280ps7.

ORIGINAL ARTICLE

Success rate of thyroid remnant ablation for differentiated thyroid cancer based on 5550 MBq post-therapy scan

I. Hommel[1]*, G.F. Pieters[1], A.J.M. Rijnders[2], M.M. van Borren[3], H. de Boer[1]

Departments of [1]Internal Medicine, [2]Nuclear Medicine, [3]Clinical Chemistry, Rijnstate Hospital, Arnhem, the Netherlands, *corresponding author: tel.: +31 (0)6-16962998, email: i.hommel@gelre.nl

ABSTRACT

Background: Success rate of thyroid remnant ablation in patients with low-risk differentiated thyroid cancer (DTC) is commonly based on measurement of serum thyroglobulin levels and 185 MBq (5 mCi) diagnostic ^{131}I scanning or neck ultrasound, performed 6-9 months after ablation. In the present study, we report the rates of successful ^{131}I ablation based on a 5550 MBq (150 mCi) post-therapy scan performed 6-9 months after ablation.

Methods: Retrospective cohort study of 77 adult patients with DTC, stage T1-T3, N0 or N1, M0, demonstrating thyroid remnant uptake one week after a 2775 MBq (75 mCi) ablation dose. Six to nine months later, all patients received a 5550 MBq dose of ^{131}I, followed by a post-therapy scan after one week. Complete thyroid ablation was defined as no thyroid remnant uptake and a thyroglobulin level < 0.2 μg/l after thyroid hormone withdrawal.

Results: Thyroid ablation was complete in 20 patients (26%). Forty-eight patients (62%) demonstrated persistent remnant uptake. This was associated with thyroglobulin levels > 0.2 μg/l in 24/48, and positive thyroglobulin antibodies in 4/48 patients.

Conclusion: Thyroid remnant ablation success assessed by 5550 MBq post-therapy scanning was much lower than reported in studies evaluating ablation success based on 185 MBq diagnostic ^{131}I scanning or neck ultrasound. The latter techniques may be too inaccurate to detect thyroid remnants and thus may not be sufficiently reliable to predict long-term disease outcome.

KEYWORDS

Differentiated thyroid carcinoma, iodine ablation dose

INTRODUCTION

In most centres, standard treatment for differentiated thyroid carcinoma (DTC) comprises total thyroidectomy followed by radioiodine (^{131}I) ablation. In low-risk patients the aim of ^{131}I ablation is to destroy thyroid remnants and local micro-metastases, and to facilitate the use of serum thyroglobulin measurements to detect disease recurrence.[1,2] In high-risk patients complete thyroid ablation is needed to enable optimal ^{131}I uptake in metastases.

The optimal dose of radioiodine that is required for successful remnant ablation is under debate.[3] Historically, ablation doses have varied from 740-5550 MBq (20-150 mCi).[4] Recently it has been suggested that thyroid ablation using low-dose ^{131}I (1110 MBq, 30 mCi) may be equally successful as high-dose treatment (3700 MBq, 100 mCi) in patients with DTC tumour stage T1-T3 without evidence of microscopic residual disease or distant metastasis.[5,6] These findings are likely to have a major impact on the treatment of DTC worldwide. It coincides with the gradually increasing concern for long-term adverse events of ^{131}I treatment, and the wish to reduce radiation exposure and shorten hospital isolation time. The Dutch guideline on treatment of DTC has already been adjusted based on the results of these studies. Recognising that 91% of the patients included in these studies had T1-T2 tumours, this guideline now recommends the 1110 MBq dose for thyroid remnant ablation for patients with T1-T2 DTC.[5-7]

The criteria to define successful ablation in the two landmark studies referred to above were thyroid-stimulating hormone (TSH)-stimulated thyroglobulin levels less than 1.0 and 2.0 μg/l, respectively (i.e. about 10 times the assay's detection level), combined with a negative

^{131}I diagnostic whole body scan (185 Mbq, 5 mCi) and/or negative neck ultrasound 6-9 months after ablation.[5,6] However, it is uncertain whether these surrogate short-term markers can reliably predict differences in long-term outcome. Accurate prediction of long-term outcome is a major challenge in low-risk DTC because recurrent disease develops very slowly. About 10-15 years of follow-up are required to obtain a well-founded assessment of treatment efficacy.[8]

As residual disease not detected at completion of initial treatment is the main risk factor for 'disease recurrence' it will be important to employ the most sensitive technique available for the detection of thyroid remnants. In the present study, we report the results of 2775 MBq (75 mCi) ^{131}I ablation, using the combination of an undetectable thyroglobulin after thyroid hormone withdrawal and a negative post-therapy whole body scan as criteria for successful thyroid ablation.

MATERIALS AND METHODS

Study design, setting and patients

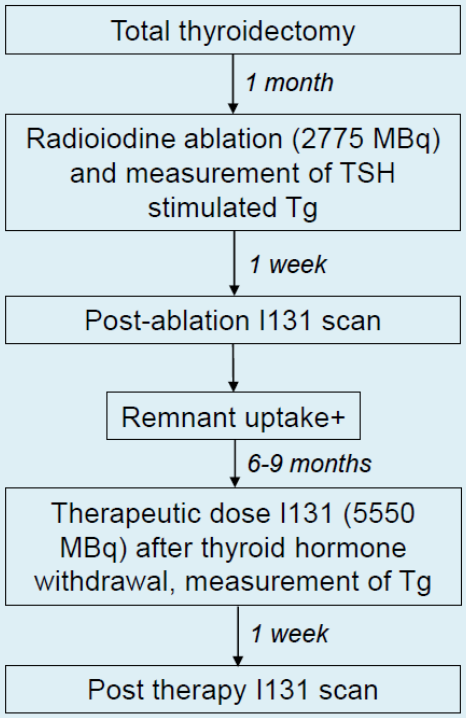
This is a retrospective observational cohort study performed in a tertiary teaching hospital. We included adult patients (>18 years) treated for DTC in the period 1998-2012. The inclusion criteria were:

(1) histologically confirmed papillary or follicular thyroid carcinoma, without evidence of microscopic residual disease;
(2) TNM stage T1-T3, N0 or N1, and no evidence of distant metastases (M0);
(3) Ablation dose 2775 MBq; and
(4) Presence of thyroid remnant uptake on the post-ablation ^{131}I whole body scan one week after ablation.

Treatment

The treatment schedule used in this study is summarised in *figure 1*. Thyroidectomy was performed by dedicated thyroid surgeons with a personal thyroid surgery volume of at least 30, predominantly benign, cases per year. A total of four different surgeons were involved during the study period. Pre-ablation scintigraphy for assessment of the percentage neck retention was discontinued in 2000 because previous measurements had repeatedly demonstrated retentions < 5%. Thyroid hormone replacement was withheld until ^{131}I ablation, which was performed four weeks after surgery. All patients received a fixed 2775 MBq (75 mCi) ^{131}I ablation dose. This has been the standard treatment for T1-T3 DTC in our centre for the past 20 years. Whole body scanning was performed one week after ablation. All patients demonstrating thyroid remnant uptake on the post-ablation scan received a 5550 MBq therapeutic dose of ^{131}I, 6-9 months later, after thyroid hormone withdrawal for four weeks. Patients were instructed to take a low iodine diet for one week before both ablation and therapy.

Figure 1. *Treatment schedule for the patients included in this study*



Outcome measures

Outcome measures were TSH-stimulated serum thyroglobulin levels and whole body scanning, performed one week after the therapeutic dose of ^{131}I. However, in patients with anti-thyroglobulin antibodies, ablation success was based on the results of the post-therapy scan only. Thyroid bed uptake after ^{131}I therapy was assessed by visual inspection of planar images. Single photon emission computed tomography or threshold measurements were not employed. Serum levels of TSH, TSH-stimulated thyroglobulin and anti-thyroglobulin antibodies were measured just before ablation or the administration of the therapeutic dose of ^{131}I.

Complete thyroid ablation was defined as (1) no visible uptake in the thyroid bed on the post-therapy scan despite a serum TSH level > 30 mU/l, and (2) a TSH-stimulated thyroglobulin level below the assay's detection limit (< 1.0 µg/l before 2004, and < 0.2 µg/l thereafter) and (3) absence of thyroglobulin antibodies.

Laboratory assays

TSH was measured by electrochemiluminescence assay (Modular E170, Roche). Thyroglobulin and anti-thyroglobulin levels were measured by immunochemiluminometric assay (Immulite 2000 XPI, Siemens). Up to 2004 the thyroglobulin detection limit was 1.0 µg/l, thereafter it declined to 0.2 µg/l. The anti-thyroglobulin assay has an analytical sensitivity of 2.2 IU/ml. Anti-thyroglobulin levels exceeding 40 IU/ml were considered positive for the presence of thyroglobulin antibodies.

Statistical analysis

Results are shown as median values and range. Descriptive statistics were used to summarise the data. A paired t-test was used to evaluate the change in thyroglobulin levels after ablation. A p-value < 0.05 was considered to be statistically significant.

RESULTS

Between 1998 and 2012, 111 patients underwent total thyroidectomy and ^{131}I thyroid remnant ablation for DTC (*figure 2*). Thirty-four patients did not meet the inclusion criteria that were set for this study. The reasons for exclusion were: post-ablation or post-therapy scan images not available (n = 15), tumour stage T4 (n = 9), distant metastases (n = 3), ablation dose < 2775 MBq (n = 4), and no remnant uptake after ^{131}I ablation (n = 3). The baseline characteristics of the 77 patients included in the study are summarised in *table 1*. Of the patients, 83% had papillary carcinoma and 17% had follicular carcinomas. Patients with mixed papillary/follicular carcinoma were classified as papillary carcinoma, and those with Hurthle cell carcinoma were classified as follicular carcinoma. DTC was multifocal in four patients (5%); 81% had T1-T2 tumours, including two patients with a T1a tumour. The remaining patients had T3 tumours (19%). Microscopic examination indicated that carcinoma resection had been radical in all patients. Neck lymph node involvement, either based on tissue examination or by post-ablation ^{131}I uptake outside the thyroid bed, was diagnosed in 10% of patients. Lymph node involvement was found in 10% of T1 tumours, in 16% of T2 tumours and in 0% of the patients with T3 tumours. Anti-thyroglobulin antibodies were present in 13/77 (17%) of patients. Thyroglobulin levels at ablation were not available in 8/64 anti-thyroglobulin negative patients. Six months later the number of patients with anti-thyroglobulin antibodies had declined to 6 (8%), and the number of anti-thyroglobulin negative patients with missing thyroglobulin-off measurements had decreased to 4/71 (6%).

Figure 2. *Patient inclusion*

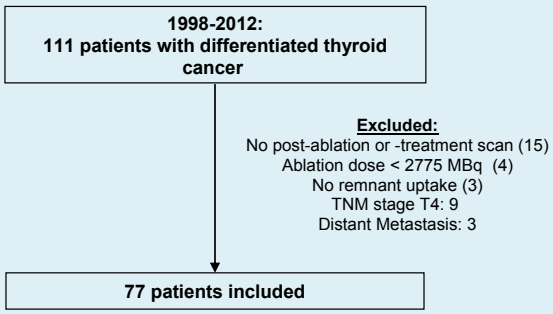


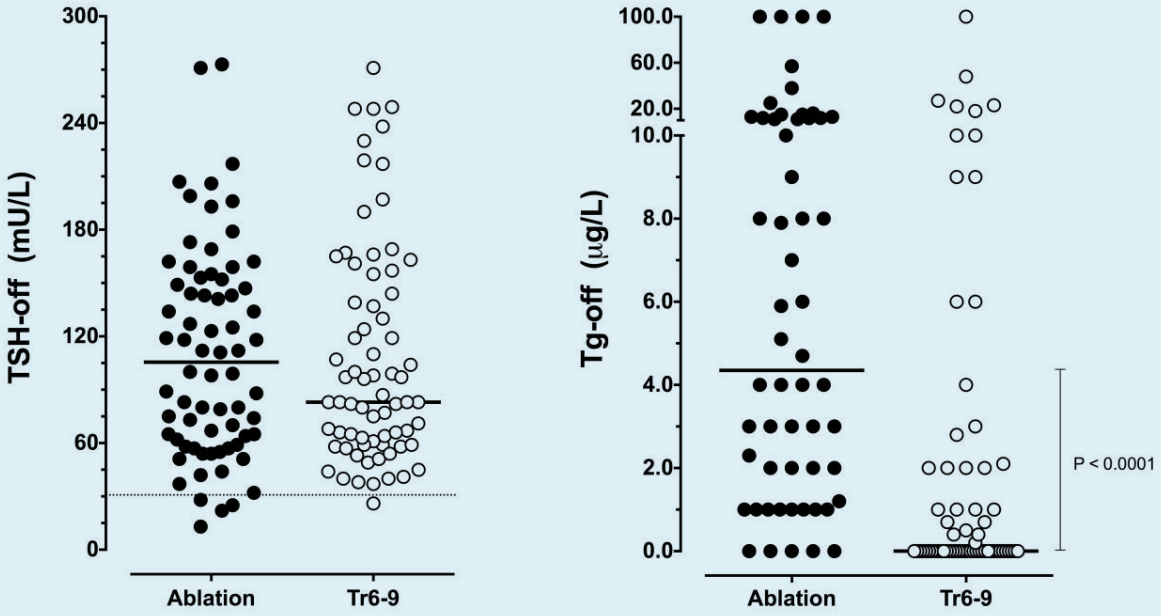
Table 1. *Baseline characteristics of 77 patients with differentiated thyroid carcinoma included in the study*

Age (years)	44 (18-85)
Male-to-female ratio	19/58
Tumour stage	
T1	40.3%
T2	40.3%
T3	19.4%
Nodal stage	
N0	89.9%
N1	10.1%
Histological findings	
Papillary	83.1%
Follicular	16.9%
Lab results at ablation	
TSH-off < 30 mU/l Median Tg-off	5.2% 4.4 µg/l
Tg-off < 2.0 µg/L	25%
Tg-antibodies present	17%

TSH = thyroid stimulating hormone; Tg = thyroglobulin.

The TSH response to a four-week thyroid hormone withdrawal varied considerably with levels ranging from 13-273 mU/l (*figure 3*). At ablation, four patients had a serum TSH < 30 mU/l. Nevertheless, they all demonstrated thyroid remnant uptake on the post-ablation scan and were therefore included in the evaluation. A post-withdrawal TSH < 30 mU/l occurred in only one patient at treatment 6-9 months later (Tr_{6-9}). In patients without anti-thyroglobulin antibodies, thyroglobulin-off was detectable in 51/64 (80%) patients at the time

Figure 3. *Serum TSH and thyroglobulin levels after a four-week withdrawal of thyroid hormone (TSH-off and thyroglobulin-off, respectively), at ablation and at ^{131}I-therapy 6-9 months later (Tr_{6-9}). Median levels are shown as solid lines*



of ablation and in 30/71 (42%) at Tr_{6-9} (*figure 3*). At Tr_{6-9} thyroglobulin-off levels ≥ 2.0 µg/l were present in 19/71 (27%) anti-thyroglobulin negative patients. The post-treatment scan at Tr_{6-9} was negative for thyroid remnant uptake in 29/77 patients (38%). In 20 of 77 patients (26%) thyroid ablation was complete, as demonstrated by a negative scan and stimulated thyroglobulin levels below the detection limit. Of the remaining nine patients with a negative scan, six had detectable thyroglobulin levels (range 0.7-27 µg/l), two had anti-thyroglobulin antibodies and a thyroglobulin value was missing in one patient. Persistent thyroid remnant uptake was demonstrated in 48 of 77 patients (62%). In 24 patients this was associated with thyroglobulin levels above the detection limit, four patients had anti-thyroglobulin antibodies, and thyroglobulin values were missing in three subjects. Of note, 35% of the patients with a positive scan (17/48) had a stimulated thyroglobulin below the detection limit. This occurred in 8/39 patients treated before 2004 when the thyroglobulin detection limit was 1.0 µg/l, and in 9/38 patients treated between 2004-2012 who were tested against a thyroglobulin detection limit of 0.2 µg/l.

DISCUSSION

This study demonstrates that a 2775 MBq ^{131}I ablation dose is associated with complete ablation in only 26% of patients if judged by very rigorous criteria. Complete ablation increased to 38% if defined as negative thyroid bed uptake, and to 55% if defined as an undetectable serum thyroglobulin upon thyroid hormone withdrawal in anti-thyroglobulin negative patients. All these figures are much lower than the recently reported success rates of 90% after an ablation dose as low as 1110 Mbq.[5,6] Several factors may explain the low success rate we observed with an ablation dose that was 2.5 times higher than used in these two comparator studies. More extensive pre-surgical disease may have contributed but a major impact is considered unlikely. In Mallick's study 22% of patients had T3 tumours, whereas patients with T3 tumours were excluded in Schlumberger's study. Despite this marked difference in tumour load, the ablation success was comparable for both studies, suggesting that pre-surgery tumour size was not an important determinant of ablation success. This is also supported by our observation that after the exclusion of the patients with T3 tumours, ablation success only increased from 26 to 27%. In contrast, post-surgery remnant thyroid tissue volume, judged by median post-surgical thyroglobulin levels at ablation, may have had an impact. The median postoperative thyroglobulin level of 4.4 µg/l in our population was nearly twice as high as that observed in Mallick's study and this could have reduced our ablation success rate to some extent. However, the marked difference in outcome is probably best explained by differences in the definition

of ablation success. In the present study this was defined as a serum thyroglobulin below the assay detection level and no visible uptake in the thyroid bed on a post-therapy scan performed one week after a 5550 Mbq dose of ^{131}I, administered 6-9 months after ablation. This is currently the most rigorous definition of ablation success. In contrast, the two comparator studies used a more lenient definition of ablation success. They employed a thyroglobulin cut-off of 1.0 and 2.0 µg/l, respectively, i.e. cut-offs that were well above their assay's functional detection limits and 5-10 times higher than in our study, thus reducing the sensitivity to detect remnant thyroid tissue and increasing the chance of an ablation being incorrectly classified as successful. Nevertheless, the impact of these differences in thyroglobulin cut-offs is probably small. When we used a thyroglobulin cut-off of 2.0 µg/l, ablation success barely increased, i.e. from 26% to 29%. This strongly suggests that differences in imaging technique, in particular the difference in sensitivity of a 185 MBq diagnostic whole body scan as compared with a 5550 MBq post-therapy scan, is the most important factor explaining the differences in ablation success. Obviously, a more sensitive technique to detect thyroid remnants will result in lower success rates.

If complete remnant ablation is crucial to achieve the lowest possible recurrence of disease, then the most rigorous criteria for successful ablation should be employed to avoid incorrect classifications of complete remission. This implies that efficacy and safety of new approaches such as lowering of the ablation dose, the decision to discharge low-risk patients from outpatient monitoring five years after achievement of complete remission, etc., should be tested against these strict criteria. More lenient criteria will increase short-term success rates but are likely to increase long-term recurrence rates. With these considerations in mind, and the observation that a 2775 MBq ablation dose fails to achieve complete remnant ablation in at least two-thirds of the patients, it appears unwise to lower the ablation dose to 1110 MBq without evidence of the long-term consequences of such a policy, or without concomitant recommendations of follow-up strategies that guarantee early recurrence detection and effective (rescue) treatment schedules. Disproportionate concern about the potential risk of second primary malignancy after ^{131}I treatment should not dominate the discussion about the most appropriate ablation dose. In a recent study of nearly 2500 patients with DTC, this risk only increased after a cumulative dose of 37,000 MBq (1000 mCi).[9] At present, the evidence favouring either high- or low-dose ablation remains inconclusive. Low-dose treatment and lenient criteria of ablation success may give rise to an increasing number of patients with serum thyroglobulin levels exceeding the limit of detection during follow-up. This may create diagnostic and therapeutic dilemmas that are yet unresolved and cause an increase in patient anxiety. For example, definitive answers to questions such as which thyroglobulin cut-off levels are safe to distinguish normal residual thyroid tissue from persistent thyroid cancer are still under debate.[10] On the other hand, high-dose ablation and follow-up with very strict criteria for cure are likely to be associated with a higher number of ^{131}I treatments, adverse events related to episodes of thyroid hormone withdrawal, longer time spent in hospital and a higher total cumulative dose of ^{131}I, and might be considered to be overtreatment. As long as the benefit-to-harm ratio of both approaches is not known, we recommend caution in the follow-up of patients treated with low-dose ablation.

In conclusion, most patients in this study had persistent thyroid remnants despite ^{131}I ablation with a dose that was 2.5 times higher than the dose recently recommended for the ablation of T1-T2 DTC. This observation may serve to illustrate that careful monitoring is warranted after low-dose ablation therapy, and that follow-up guidelines may need to be redefined to optimise detection of disease recurrence and to adjust the specifications of the indications of treatment. Long-term studies comparing the results of low- and high-dose ablation will be required to further clarify the impact of the 1110 MBq ablation dose on disease recurrence rates and mortality.

DISCLOSURES

All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

REFERENCES

1. Mazzaferri EL, Kloos RT. Clinical review 128: current approaches to primary therapy for papillary and follicular thyroid cancer. J Clin Endocrinol Metab. 2001;86:1447-63.
2. Tsang RW, Brierley JD, Simpson WJ, Panzarella T, Gospodarowicz MK, Sutcliffe SB. The effects of surgery, radioiodine, and external radiation therapy on the clinical outcome of patients with differentiated thyroid carcinoma. Cancer. 1998;82:375-88.
3. Links TP, van der Horst-Schrivers AN. Thyroid cancer: successful remnant ablation-what is success? Nat Rev Endocrinol. 2012;8:514-5.

4. Cheng W, Ma C, Fu H, et al. Low- or high-dose radioiodine remnant ablation for differentiated thyroid carcinoma: a meta-analysis. Clin Endocrinol Metab. 2013;98:1353-60.

5. Mallick U, Harmer C, Yap B, et al. Ablation with low-dose radioiodine and 3238 thyrotropin alfa in thyroid cancer. N Engl J Med. 2012;366:1674-85.

6. Schlumberger M, Catargi B, Borget I, et al. Strategies of radioiodine 3245 ablation in patients with low-risk thyroid cancer. N Engl J Med. 2012;366:1663-73.

7. Guideline for the diagnosis, treatment and follow-up of patients with differentiated (non-medullary) thyroid carcinoma. Dutch Endocrine Society Dutch and Comprehensive Cancer Centre the Netherlands. 2015. http://www.oncoline.nl/schildkliercarcinoom.

8. Verburg FA, Mäder U, Reiners C, Hänscheid H. Long-term survival in differentiated thyroid cancer is worse after low-activity initial post-surgical I-131 therapy in both high- and low-risk patients. J Clin Endocrinol Metab. 2014;99:4487-96.

9. Khang AR, Cho SW, Choi HS, et al. The risk of second primary malignancy is increased in differentiated thyroid cancer patients with a cumulative (131)I dose over 37 GBq. Clin Endocrinol. 2015;83:117-23.

10. American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer, Thyroid 2015, online ahead of editing October 14.

ORIGINAL ARTICLE

The influence of oral contraceptives on overnight 1 mg dexamethasone suppression test

M. Vastbinder[1], M. Kuindersma[2]*, A.H. Mulder[2], M.P Schuijt[3], A.H. Mudde[2]

The first two authors contributed equally

[1]Department of Internal Oncology, Erasmus MC, Rotterdam, the Netherlands, [2]Department of Internal Medicine, Slingeland Hospital, Doetinchem, the Netherlands, [3]Department of Clinical Chemistry and Laboratory Medicine, Slingeland Hospital, Doetinchem, the Netherlands, *corresponding author: tel.: +31 (0)314 - 329 659, fax: +31 (0)314 - 329 150, email: a.mudde@slingeland.nl

ABSTRACT

Background: In suspected hypercortisolism, the 1 mg dexamethasone suppression test is the usual initial test. In fertile women, false-positive test results are often due to the use of oral contraceptives. By elevating cortisol-binding globulin these contraceptives increase the total serum cortisol concentration. The aim of this study was to assess the duration and degree of influence of oral contraceptives on the low-dose dexamethasone suppression test.

Methods: Thirteen healthy female volunteers without symptoms or signs of overt hypercortisolism, aged 18-55 years, who were using oral contraceptives, underwent a 1 mg dexamethasone suppression test. Tests were repeated one and six weeks after withdrawal of the contraceptive. In addition, 24-hour urinary cortisol excretion and late-night salivary cortisol were measured.

Results: Of the 13 volunteers (62%) eight had inadequate suppression of cortisol by 1 mg dexamethasone while using oral contraceptives. One week after the contraceptive was withdrawn, the number of false-positive results significantly decreased to 1 (8%, $p < 0.02$). Six weeks after discontinuation, all tests were normal. None of the 24-hour urinary cortisol samples and just one late-night salivary cortisol level was elevated.

Conclusion: The results of the 1 mg dexamethasone suppression test performed one week after cessation of oral contraceptives are accurate in almost all subjects. In case of inadequate suppression, a second test may be performed after six weeks. In this manner the 1 mg dexamethasone suppression test can reliably be done at the end of a seven-day break from contraceptive use in nearly all cases.

KEYWORDS

Cushing syndrome/diagnosis, dexamethasone, oral contraceptives

INTRODUCTION

Testing for Cushing's syndrome is common daily practice, since many of the signs and symptoms of this syndrome are nonspecific and common in the general population. Moreover, the increasing prevalence of adrenal incidentaloma has augmented the need for testing to exclude Cushing's syndrome.[1] Therefore the 1 mg dexamethasone suppression test (DST), as a first-line screening test in the work up for Cushing's syndrome, is commonly used in clinical practice. In this test serum total cortisol concentration is measured after oral administration of 1 mg of dexamethasone. At a cut-off value of 50 nmol/l the DST has a high sensitivity (95%) and a reasonable specificity (80%).[2-4]

Test results may be influenced by the use of drugs altering the clearance of cortisol, such as CYP3A4 inhibitors and inducers, or causing an elevation of protein-bound cortisol fraction. In serum, cortisol is bound to albumin and cortisol-binding globulin. A small fraction is unbound and metabolically active.[5] Oral contraceptives increase the total cortisol concentration by increasing circulating cortisol-binding globulin, whereas the unbound fraction is not influenced.[6] Thus, in women using oral contraceptives, measurement of total cortisol in a 1 mg DST could be elevated although no increased activity of cortisol is present.

As a consequence, false-positive test results are found in up to 50% in women using oral contraceptives.[2] In clinical practice, oral contraceptives are usually discontinued for six weeks in order to prevent false-positive results of the DST. However, evidence for this practice is lacking. Previously it was suggested that serum cortisol 16 hours after ingestion of 1 mg of dexamethasone is not influenced by oral contraceptives.[7,8] Another suggestion to overcome the influence of oral contraceptives in the DST was to perform the test with double-dosed dexamethasone.[9] We performed a study to determine if a 1 mg DST can be reliably performed at the end of a seven-day break from oral contraceptives. While urinary and salivary cortisol is not bounded to cortisol-binding globulin and therefore supposed not to be influenced by oral contraceptives, we also determined cortisol in late-night sampled saliva and in 24-hour urine samples.[10-12]

PATIENTS AND METHODS

The study was approved by the Central Committee on Research Involving Human Subjects (CCMO) under number NL31515. Fifteen healthy female volunteers, aged 18-55 years old, who were using oral contraceptives containing at least 30 µg of ethinyl oestradiol were recruited. Exclusion criteria were Cushing's syndrome, or using medication that influences cortisol metabolism (e.g. phenytoin, barbiturates, ephedrine, rifampicin). Pregnancy was excluded in all patients by a pregnancy test. Two subjects withdrew from the study for personal reasons. Here we present the results of the 13 subjects who completed the study.

Standard 1 mg DSTs were performed in all participants. These volunteers were given 1 mg of dexamethasone orally at 23.00 hours. The next day blood samples were taken between 8.00 and 9.00 hours, measuring fasting total serum cortisol using an electrochemiluminescence immunoassay (ECLIA) on the Cobas e601 platform (reference < 50 nmol/l). One day before the DST was performed, participants collected a late-night saliva between 23.00 and 24.00 hours using a salivette with a polyester swab (Sarstedt) and a 24-hour urinary sample. Cortisol concentrations in saliva and urine were measured with an ECLIA, reference values < 15 nmol/l and < 0.38 µmol/24 hours, respectively.[13] Prior to analysis, the urine samples were prepared with a dichloromethane extraction procedure. All tests were performed while oral contraceptives were being used and repeated one and six weeks after discontinuation of the oral contraceptives.

The primary endpoint was the level of fasting cortisol in the 1 mg DST after discontinuation of the oral contraceptive. The differences in cortisol levels were compared pair-wise with each other (longitudinally). Secondary endpoints were the longitudinal changes in late-night salivary cortisol and 24-hour urine cortisol.

DATA ANALYSIS

Analysis was performed using GraphPad Prism (GraphPad Software Inc., version 5). Results of the 1 mg DST with and without oral contraceptives were compared using Fisher's exact test for categorical data. The Kruskal-Wallis test and a Dunn's multiple comparison test were used to determine differences in median serum, urinary and salivary cortisol results.

RESULTS

Eight of thirteen volunteers had inadequate suppression in the 1 mg DST while using contraceptives ('positive test'). One week after cessation of the oral contraceptive, the number of positive tests decreased to one, excluding Cushing's syndrome in 38% and 92% respectively ($p < 0.02$). After six weeks, none of the 1 mg DSTs were positive (*figure 1*).

The 24-hour urinary cortisol excretion was not significantly elevated in any of the participants. All results of the late-night salivary cortisol measurements were below the cut-off value, except for one sample (40 nmol/l) (*figure 2*). Hypercortisolism was not established in any of the participants.

DISCUSSION

This study shows that the 1 mg DST can be performed reliably one week after cessation of oral contraceptives.

Figure 1. *Results of DST. One line represents one volunteer*

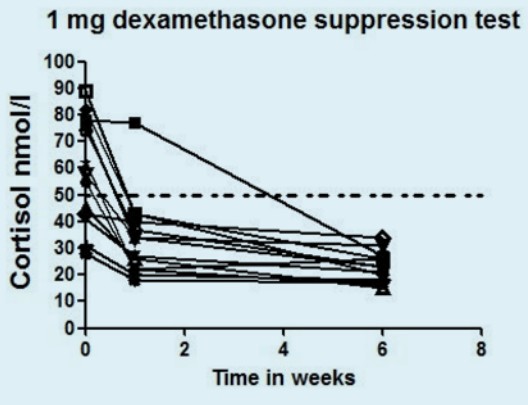
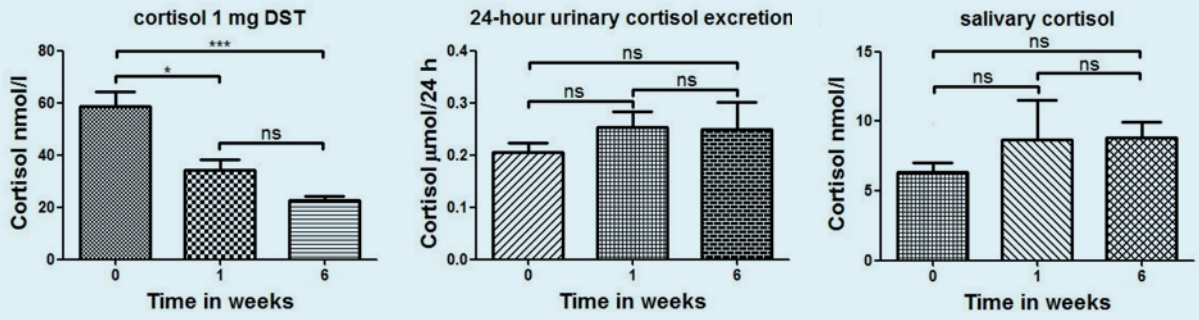


Figure 2. *Median ± IQR serum cortisol after 1 mg dexamethasone with a significant decrease after 1 and 6 weeks. Median ± IQR urinary and salivary cortisol respectively showed no significant results*



ns = not significant, *p < 0.05, ***p < 0.001.

In our study we were able to exclude hypercortisolism in 12 of 13 (92%) volunteers one week after cessation of oral contraceptives. The one subject with insufficient cortisol suppression was using an oral contraceptive containing 30 mg ethinyl oestradiol, comparable with the dosage used in the other subjects in our study. None of the 24-hour urinary cortisol samples and just one late-night salivary cortisol level were elevated, indicating that the DST was correctly ruling out Cushing's disease. The reason for the outlier in salivary cortisol may be minor intra-oral bleeding, although the cause in this specific case was unclarified. The subject was different from the one with inadequate cortisol suppression in the DST.

Our data endorse the previous literature where the influence of ethinyl oestradiol on total cortisol was described. This is caused by elevated cortisol-binding globulin levels without an increase in the fraction of free cortisol, the biologically active cortisol. Since total cortisol is measured in the standard 1 mg DST, oral contraceptives containing ethinyl oestradiol make the DST unreliable. This relationship has been demonstrated before, although the duration of the influence of oral contraceptives on cortisol-binding globulin, and thereby on the 1 mg DST, remained unclear.[6-8]

This is in contradiction with previous literature in which the effect of oral contraceptives on the DST was considered very limited or nearly absent.[7,8] The strength of our study is that we repeated the DST in the same subjects, thereby demonstrating that this test is actually more often false-positive while using oral contraceptives and normalises when these drugs are discontinued.

In clinical practice, oral contraceptives are often discontinued for six weeks before a DST is performed, to rule out any influence of ethinyl oestradiol on the test results. However, this practice is patient unfriendly and may delay the diagnostic workup. Given the high diagnostic yield after just one week of cessation of contraceptives in our study, we propose a different procedure in which the dexamethasone suppression test is performed one week after withdrawing the contraceptive. In case of inadequate suppression, a second suppression test can be performed after the usual period of six weeks of withdrawal.

In conclusion, in this study we have shown that a 1 mg DST can reliably be performed one week after withdrawal of contraceptives containing ethinyl oestradiol. In case of inadequate suppression, a second suppression test can be performed after the usual period of discontinuation of six weeks. Alternatively, salivary and urinary cortisol can be measured, since oral contraceptives do not influence these measurements.[10-12] The accelerated performance of the DST does not cause a loss of the diagnostic value of the test, and is much more patient friendly, because women do not have to extend their discontinuation of use of oral contraception to more than one week. Moreover, the reliability of the contraception is not disturbed when DST is performed at the end of a seven-day break of oral contraception usage.

ACKNOWLEDGEMENTS

Prof. A.R.M.M. Hermus MD PhD, internist-endocrinologist, Radboud UMC, Nijmegen and Prof. M. den Heijer MD PhD, internist-endocrinologist, VUMC, Amsterdam, for their critical support.

DISCLOSURES

The authors declare no conflict of interest. No funding or financial support was received.

REFERENCES

1. Longo DL, Harrison TR. Harrison's principles of internal medicine. 18th ed. Melmed S JJ, editor. New York: McGraw-Hill; 2012.
2. Nieman LK, Biller BM, Findling JW, et al. The diagnosis of Cushing's syndrome: an Endocrine Society Clinical Practice Guideline. J Clin Endocrinol Metab. 2008;93:1526-40.
3. Elamin MB, Murad MH, Mullan R, et al. Accuracy of diagnostic tests for Cushing's syndrome: a systematic review and metaanalyses. J Clin Endocrinol Metab. 2008;93:1553-62.
4. Wood PJ, Barth JH, Freedman DB, Perry L, Sheridan B. Evidence for the low dose dexamethasone suppression test to screen for Cushing's syndrome--recommendations for a protocol for biochemistry laboratories. Ann Clin Biochem. 1997;34:222-9.
5. Vermeulen A. Steroïden en bindende eiwitten. Ned Tijdschr Klin Chem. 2003;28:174-6.
6. Qureshi AC, Bahri A, Breen LA, et al. The influence of the route of oestrogen administration on serum levels of cortisol-binding globulin and total cortisol. Clin Endocrinol (Oxf). 2007;66:632-5.
7. D'haenen H, de Weert D, Ansseau M. Lack of effects of hospitalization and oral contraceptives on DST results in control subjects. Biol Psychiatry. 1987;22:1499-502.
8. Ansseau M, Leboulle D, Sulon J, von Frenckell R, Legros JJ. Oral contraceptives and the dexamethasone suppression test. Psychoneuroendocrinology. 1993;18:37-43.
9. Nickelsen T, Lissner W, Schöffling K. The dexamethasone suppression test and long-term contraceptive treatment: measurement of ACTH or salivary cortisol does not improve the reliability of the test. Exp Clin Endocrinol. 1989;94:275-80.
10. Manetti L, Rossi G, Grasso L, et al. Usefulness of salivary cortisol in the diagnosis of hypercortisolism: comparison with serum and urinary cortisol. Eur J Endocrinol. 2013;168:315-21.
11. Inder WJ, Dimeski G, Russell A. Measurement of salivary cortisol in 2012 - laboratory techniques and clinical indications. Clin Endocrinol (Oxf). 2012;77:645-51.
12. Liening SH, Stanton SJ, Saini EK, Schultheiss OC. Salivary testosterone, cortisol, and progesterone: two-week stability, interhormone correlations, and effects of time of day, menstrual cycle, and oral contraceptive use on steroid hormone levels. Physiol Behav. 2010;99:8-16.
13. Yaneva M, Kirilov G, Zacharieva S. Midnight salivary cortisol, measured by highly sensitive electrochemiluminescence immunoassay, for the diagnosis of Cushing's syndrome. Cent Eur J Med. 2009;4:59-64.

ORIGINAL ARTICLE

Guideline-related barriers to optimal prescription of oral anticoagulants in primary care

A.L. Beukenhorst[1]*, D.L. Arts[2], W. Lucassen[2], K.J. Jager[1], S.N. van der Veer[3]

Departments of [1]Medical Informatics, [2]Family Medicine, Academic Medical Center, Amsterdam, the Netherlands, [3]European Renal Best Practice (ERBP) Methods Support Team, University Hospital Ghent, Belgium, Health e-Research Centre, Institute of Population Health, University of Manchester, Manchester, United Kingdom, *corresponding author: tel.: +31 (0)20 - 566 52 69, email: a.beukenhorst@gmail.com

ABSTRACT

Guidelines provide recommendations for antithrombotic treatment to prevent stroke in people with atrial fibrillation, but oral anticoagulant prescriptions in Dutch primary care are often discordant with these recommendations. Suboptimal guideline features (i.e. format and content) have been suggested as a potential explanatory factor for this type of discordance. Therefore, we systematically appraised features of the Dutch general practitioners' (NHG) atrial fibrillation guideline to identify guideline-related barriers that may hamper its use in practice. We appraised the guideline's methodological rigour and transparency using the Appraisal of Guidelines, Research and Evaluation (AGREE) II tool. Additionally, we used the Guideline Implementability Appraisal (GLIA) tool to assess the key recommendations on oral anticoagulant prescription. The editorial independence of the guideline group scored highly (88%); scores for other aspects of the guideline's methodological quality were acceptable, ranging from 53% for stakeholder involvement to 67% for clarity of presentation. At the recommendation level, the main implementation obstacles were lack of explicit statements on the quality of underlying evidence, lack of clarity around the strength of recommendations, and the use of ambiguous terms which may hamper operationalisation in electronic systems. Based on our findings we suggest extending stakeholder involvement in the guideline development process, standardising the layout and language of key recommendations, providing monitoring criteria, and preparing electronic implementation parallel with guideline development. We expect this to contribute to optimising the NHG atrial fibrillation guideline, facilitating its implementation in practice, and ultimately to improving antithrombotic treatment and stroke prevention in people with atrial fibrillation.

KEYWORDS

Atrial fibrillation, anti-thrombotic treatment, cardiology/standards, practice guidelines as topic, primary stroke prevention

INTRODUCTION

Prevalence of atrial fibrillation (AF), a common arrhythmia, has increased over the last 30 years.[1] In 2006 its prevalence ranged from 0.7% (age group 55-59) to 17.8% (age group > 85). AF increases the risk and severity of stroke.[2,3] Antithrombotic therapy with oral anticoagulation (OAC) decreases this stroke risk, but at the same time increases the risk of major bleeding.[2] National and international clinical practice guidelines on AF management provide guidance on how to weigh these risks against expected benefits, and include recommendations on antithrombotic treatment.[3,4] Yet, antithrombotic treatment is often not in line with these recommendations.[5,6] For example, a study in the Netherlands estimated that less than 50% of AF patients received treatment according to national guidelines.[6]

Various types of barriers may thwart physicians in following guidelines in clinical practice, such as lack of familiarity with the guideline's content, lack of skills or resources to change current practice or patients not reconciling with guideline recommendations.[7] Suboptimal

guideline features (i.e. format and content) may also hamper implementation.[7,8] Lugtenberg et al. reported this as one of the barriers that hindered general practitioners (GPs) in following the Dutch College of General Practitioners (NHG) AF guideline.[9]

Improving these features may positively affect guideline use.[8,10,11] Therefore, this study systematically appraised the format and content of the NHG AF guideline[4] to identify suboptimal features that may hamper its use in Dutch primary care. We focused on the guideline section related to prescription of OACs for stroke prevention in AF patients. The results of this appraisal may contribute to improving features of future AF guideline versions, as well as to developing tools and strategies for AF guideline implementation.

MATERIAL AND METHODS

The NHG AF guideline

The NHG aims to promote evidence-based primary care by bridging the gap between theory and practice. With 12,000 members,[12] they cover around 80% of all Dutch GPs[13] and nurse practitioners.[14] The NHG has developed over 100 guidelines covering the diagnosis and treatment of acute and chronic conditions, with ten guidelines related to cardiovascular diseases. In the current study, we reviewed the 2013-updated version of the NHG guideline on diagnosis and treatment of patients with atrial fibrillation.[4] It consists of 34 pages of background information on AF, recommendations for practice, endnotes and references. We focused on three key recommendations: (I) Eligibility criteria for OACs; (II) Which type of OAC to prescribe: coumarin derivatives versus new OACs (NOACs); and (III) Type and dosage of coumarin derivatives. The recommendations are displayed in *table 1*.

Table 1. *Key recommendations on oral anticoagulants prescription included in the NHG AF guideline*

Recommendation 1 [R1] – Eligibility criteria for OACs
The following recommendations apply to patients aged 65 years and older with atrial fibrillation; younger patients are eligible for assessment by a cardiologist:
- Recommend oral anticoagulants to all women aged 65 and older and all men 75 years and older (i.e. for patients with a CHA2DS2-Vasc score of 2 or higher)
- Discuss with male patients aged 65 to 75 years without cardiovascular comorbidities (CHA2DS2-Vasc score of 1) that the benefits of antithrombotic medication (prevention of thromboembolism) are outweighed by the disadvantages (risk of side effects, such as bleeding) and for that reason antithrombotic medication is not indicated
- Recommend aspirin when a contraindication for oral anticoagulants is present. See table with most important contraindications for antithrombotic medication

Recommendation 2 [R2] – Which type of OAC to prescribe: coumarin derivatives versus NOACs
When an indication for oral anticoagulants is present, coumarin derivatives are preferred. Consider NOACs only if all of the following conditions are met:
  - Age below 80 (arbitrary)
  - Relatively little comorbidity
  - Normal kidney function (GFR> 50 ml / min)
  - High medication adherence
- Absolute contraindications for NOACs are:
  - Patients with a mechanical artificial heart valve
  - Patients with a (currently rare) rheumatic mitral stenosis

Recommendation 3 [R3] – type and dosage of coumarin derivatives
Choice of coumarin derivative is partly determined by agreement with the local thrombosis service. In the Netherlands, short-acting acenocoumarol 1 mg and long-acting phenprocoumon 3 mg are available
- In general, the thrombosis services recommend taking the tablets once daily in the evening
- When starting a coumarin derivative, a loading dose is given for the first days according to the table with loading doses of coumarin derivatives for different patient populations
- Self-monitoring of INR may be considered for patients who find regular monitoring by the local thrombosis service burdensome
- Coumarin derivative dosing by a thrombosis service should aim for an INR between 2.0 and 3.0

AF = atrial fibrillation; CHA2DS2-Vasc score = score calculating stroke risk in AF patients; GFR = glomerular filtration rate; INR = international normalised ratio; NHG = Nederlands Huisartsgenootschap; NOACs = new oral anticoagulants; OAC = oral anticoagulant; R = recommendation.

Systematic appraisal of guideline features

To systematically appraise the features of the NHG AF guideline, we used the Appraisal of Guidelines, Research and Evaluation (AGREE) II tool[15] and the GuideLine Implementability Appraisal (GLIA) tool.[16] Both tools are publicly available, and have previously been used for guideline appraisals.[17-19]

The AGREE II tool focuses on assessing the methodological rigour and transparency with which a guideline has been developed. It contains 23 items grouped in six domains (*table 2*, first column). Each item reflects a statement that refers to the guideline as a whole (e.g., 'Key recommendations are easily identifiable'), and is scored on a seven-point Likert scale, ranging from 1 (strongly disagree) to 7 (strongly agree).[15]

To identify obstacles to implementation of the guideline's key recommendations on OAC prescription (*table 1*), we completed GLIA appraisals. GLIA consists of 21 items in eight dimensions that – in contrast to AGREE II — are scored at the level of individual recommendations (*table 3*, first column). Each item is formulated as a question (e.g., 'Is justification for the recommendation stated explicitly?') with response categories 'yes', 'no', 'not applicable', and 'unsure'. We did not assess GLIA's global dimension that appraises the guideline in its entirety and largely overlaps with the AGREE appraisal.

Data collection and analysis

Following the AGREE II and GLIA manuals,[20,21] our appraisal panel consisted of four experts, representing a mix of clinical and methodological guideline expertise: one general practitioner (WL), one expert on antithrombotic treatment and stroke prevention in AF patients (DA), and two experts on guideline development and implementation (AB, SV). Panel members first individually performed the appraisals, using the online AGREE (www.agreetrust.org) and GLIA (eGLIA; http://nutmeg.med.yale.edu/glia) tools. They also provided additional information in free text fields to explain their scores. The appraisal process was primarily informed by the guideline document itself, but when necessary, extra information was collected from: i) the NHG website; ii) the booklet on NHG guideline development procedures;[22] and iii) a structured interview with two members of the NHG AF guideline development group. The appraisal coordinator (AB) then summarised the results as input for a group discussion based on which panel members could alter their scores when they considered this appropriate (e.g., to correct for available data that were overlooked during the initial appraisal). We discussed every item for which scores differed by more than one point, and every item for which the NHG development group members provided additional information.

AGREE II domain scores were calculated by summing up the individual appraisers' scores for each item within a domain (i.e., obtained score), and then standardising this as a percentage of the possible maximum score for that domain,[20] as follows:

$$\text{Domain score in \%} = \frac{\text{(obtained score – minimum possible score)}}{\text{(maximum possible score – minimum possible score)}}$$

As a result of the consensus procedure for GLIA scores, features were categorised as optimal ('Y' in *table 3*) or suboptimal ('N' in *table 3*). Per recommendation, we calculated the percentage of suboptimal features as follows:

$$\text{Percentage of suboptimal features} = \frac{\text{\# suboptimal features of recommendation}}{\text{total \# GLIA features – not applicable features}}$$

RESULTS

AGREE II appraisal of overall guideline features

Appendix 1 presents the individual appraisers' scores before and after group discussion. The group discussion resulted in 24 out of 92 (26%) scores being changed. The main reasons for appraisers to change their scores were: correction for available data that were overlooked during the initial appraisal (8 of 24 changes; 33%); a change of opinion following clarification of other appraisers' opinion (7 of 24; 29%); correction for additional information provided by the NHG guideline development group members (6 of 24; 25%). After the group discussion, standard deviations of item scores ranged between 0 and 1.6, with the majority (75%) being 1 or lower.

Table 2 presents the final item and domain scores assigned during the AGREE appraisal of the guideline. Domain scores ranged from 52.8% for 'Stakeholder involvement' to 87.5% for 'Editorial independence'.

GLIA appraisal of key recommendation features

Table 3 displays suboptimal features at the recommendation level, which may hinder the guideline's implementation and applicability in practice. The percentage of suboptimal features for the three key recommendations on OAC prescription (R1-3) ranged from 24% to 45%. The panel considered the domains 'Effect on process of care' and 'Measurability' optimal across all recommendations.

We found that all recommendations suffered from suboptimal decidability due to ambiguous or unclear

Table 2. *AGREE II scores of the NHG AF guideline section on prescription of oral anticoagulants*

AGREE II domains, domain scores[a], and items	Item scores (SD)[b]	Illustration of suboptimal features per domain
SCOPE AND PURPOSE (63.9%)		
Overall objective of the guideline is specifically described	4.8 (1.5)	Although some health questions are included as subheadings, the guideline does not provide an easy-to-access overview of all questions covered It is unclear whether children are included in the guideline's target population
Health questions covered by the guideline are specifically described	4.0 (1.4)	
Target patient population of the guideline is specifically described	5.8 (1.0)	
STAKEHOLDER INVOLVEMENT (52.8%)		
All relevant professional groups were included in the guideline development group	4.0 (0.8)	The guideline development group did not include cardiologists, pharmacists, neurologists, or representatives of anticoagulation clinics The national patient association was asked for external review, but it is unclear if and how their suggestions were addressed in the final guideline document
Views and preferences of the target patient population were sought	2.8 (0.5)	
Target users of the guideline are clearly defined	5.8 (1.0)	
RIGOUR OF DEVELOPMENT (63.0%)		
Systematic methods were used to search for evidence	4.8 (1.0)	Although the guideline group did apply selection criteria, these were not made explicit before evidence selection[d], and not available within the guideline document For some studies, the group performed a quality appraisal, but without applying a formal tool Furthermore, recommendations lack a summary of the quality of underlying evidence that is easy to find and interpret for guideline users It is unclear how evidence was translated into recommendations, and there is no description of how the guideline group solved any disagreements arising during recommendation formulation The Netherlands Society of Cardiology did not externally review the guideline prior to publication, but most other relevant stakeholder groups did
Criteria for selecting the evidence are clearly described	2.5 (1.3)	
Strengths and limitations of the body of evidence are clearly described[c]	3.0 (1.6)	
Methods for formulating recommendations are clearly described	2.5 (1.0)	
Health benefits, side effects, and risks were considered	6.5 (0.6)	
There is an explicit link between recommendations and supporting evidence	5.5 (1.0)	
The guideline was externally reviewed by experts prior to publication	6.5 (0.6)	
Procedure for updating the guideline is provided	7.0 (0.0)	
CLARITY OF PRESENTATION (66.7%)		
Recommendations are specific and unambiguous[e]	5.5 (1.0)	Key recommendations are not easily identifiable because they do not have one specific layout or font; some, but not all are (partly) captured in boxes. Further, some text boxes hold other types of information, such as a description of guideline development procedures, or a summary of what has changed since the previous version of the guideline
Different disease management options are clearly presented	6.0 (0.8)	
Key recommendations are easily identifiable	3.5 (0.6)	
APPLICABILITY (57.3%)		
Facilitators and barriers to application of the guideline are described	4.0 (0.8)	The guideline describes several barriers and facilitators. However, most are not clearly labelled as a such, and they are scattered throughout the text instead of grouped or summarised in one section The criteria for monitoring the dose of coumarin derivatives are clearly described, but less so for other OAC types. The guideline does not provide any criteria to audit adherence or monitor impact at the GP practice level
Guideline provides advice and tools for applying recommendations in practice	4.8 (1.5)	
Potential resource implications of applying recommendations were considered	6.0 (0.0)	
The guideline presents monitoring and/or auditing criteria	3.0 (0.8)	
EDITORIAL INDEPENDENCE (87.5%)		
Views of funding body did not influence the guideline content	6.0 (0.8)	[no substantial suboptimal features with regard to the guideline's editorial independence]
Competing interests of authors were recorded and addressed	6.5 (0.6)	

AF = atrial fibrillation; GP = general practitioner; GRADE = Grading of Recommendations Assessment, Development and Evaluation; NHG = Nederlands Huisartsgenootschap; OAC = oral anticoagulant; SD = standard deviation. [a]Domain scores range between 0 and 100% and were calculated by summing up the individual appraiser's scores for each item within a domain, and then standardising this as a percentage of the possible maximum score for that domain. [b]Average item scores of four appraisers. Item scores can range between 1 and 7, with lower scores indicating less optimal features. [c]Overlaps with GLIA item 'Is quality of evidence that supports each recommendation stated explicitly?' [d]Based on data collected during structured interviews with guideline development group members. [e]Overlaps with GLIA 'Executability' and 'Decidability' domains; see *table 3*.

Table 3. *GLIA scores of the three NHG recommendations on prescription of oral anticoagulants*[a]

GLIA domains and items ↓	Scores per recommendation			Illustration of suboptimal features[b]
	R1	R2	R3	
EXECUTABILITY				
Is the recommended action stated specifically and unambiguously?	Y	N	Y	[*R2*] GPs might not execute the action '…consider NOACs…' consistently [*R2*] Type and dose of NOACs are not specified
Is sufficient detail provided to perform the recommended action?	Y	N	Y	
DECIDABILITY				
Can one consistently determine whether each condition in the recommendation was satisfied?	N	N	N	Examples of conditions that may not be consistently applied by most GPs: [*R1 – referenced table*] Contraindications, such as 'severe disturbance of liver function'; [*R2*] 'High medication adherence'; [*R3 – referenced table*] 'Relative contraindications' to treatment with warfarin and phenprocoumon
Are all reasonable combinations of conditions addressed?	Y	Y	Y	
Is the logical relationship between conditions clear?	Y	Y	Y	
VALIDITY				
Is justification for the recommendation stated explicitly?	N	Y	Y	[*R1*] Justification is stated in end notes, but not referenced as such in the recommendation For none of the recommendations, evidence quality was systematically appraised or stated
Is quality of evidence that supports each recommendation stated explicitly?[c]	N	N	N	
FLEXIBILITY				
Is the strength of each recommendation stated explicitly?	N	N	N	For all recommendations the strength is unclear due to a lack of standardised terminology or labels
Are patient characteristics specified that permit individualisation?	Y	Y	Y	
Are practice characteristics specified that permit modification?	n.a.	n.a.	Y	
EFFECT ON PROCESS OF CARE				
Can the recommendation be executed without substantial disruption in current workflow?	Y	Y	Y	n.a.
Can the recommendation be pilot tested without substantial resource commitment?	Y	Y	Y	
MEASURABILITY				
Can adherence to this recommendation be measured?	Y	Y	Y	n.a.
Can outcomes of this recommendation be measured?	Y	Y	Y	
NOVELTY & INNOVATION				
Can the recommendation be performed without acquisition of new knowledge/skills?	Y	N	Y	[*R2*] GPs need to gain knowledge on the recommended use and (side) effects of NOACs [*R2*] NOAC prescription is discouraged, while patients may expect access to the latest drugs, or explicit room for shared decision making
Is the recommendation consistent with attitudes/beliefs of the intended audience?	Y	Y	Y	
Is the recommendation consistent with patient expectations?	Y	N	Y	
COMPUTABILITY				
Are all patient data available for electronic implementation of the recommendation?	N	N	N	All recommendations have some conditions for which data are unavailable (e.g., artificial heart valve, arrangements between patient and anticoagulation clinic), or that lack specificity (see Decidability items) [*R1*] The action in statement c ('Discuss with…') may not be specific enough to allow electronic implementation
Are the recommendation's conditions defined specifically enough for electronic implementation?	N	N	N	
Is each recommended action defined specifically enough for electronic implementation?	N	Y	Y	
Is it clear by what means a recommended action can be executed in an electronic setting?	Y	Y	Y	
Number (%) of suboptimal features per recommendation	7 (35)	9 (45)	5 (24)	

GLIA = guideline implementability appraisal; GPs = general practitioners; N = suboptimal feature/GLIA score 'No'; NHG = Nederlands Huisartsgenootschap; NOACs = new oral anticoagulants; R = recommendation; Y = optimal feature/GLIA score 'Yes'; n.a. = not applicable. [a]See *table 1* for a description of each of the three key recommendations. [b]This column contains examples to illustrate why recommendation features were considered suboptimal; each example is preceded by the number of the key recommendation – for example, [*R1*] – it refers to. [c]This item overlaps with AGREE II item 'Strengths and limitations of the body of the evidence are clearly described'.

conditions for when to apply a recommendation. For example, not all GPs may agree on what could be considered 'high' (R2) levels of medication adherence, or 'relative contraindications' for prescribing (R3 - referenced table). This, together with the lack of detail on how to prescribe NOACs (i.e., suboptimal executability of R2), led the panel to score many of the computability features as problematic. This reflects the panel's expectation that operationalising the recommendations in an electronic information system may be difficult.

The strength of recommendations, and thus the degree to which they applied to all patients, was often unclear, resulting in a suboptimal flexibility score. This stemmed from a lack of standardisation of how recommendations were formulated. Terminology ranged from *'absolute contraindications* of NOACs *are*...' (R2) to 'Coumarin derivative dosing [...] *should* aim' (R3), and from '*Recommend* oral anticoagulants ...' (R1) to '... *consider* NOACs *only*...' (R2). In some cases, actions were only suggested indirectly: '...coumarin derivatives are preferred' (R2); '...a loading dose is given for the first days' (R3).

DISCUSSION

In this study we systematically appraised features of the NHG AF guideline to identify guideline-related barriers that may hamper optimal prescription of oral anticoagulants in Dutch primary care. The editorial independence of the guideline development group scored highly; scores for all other aspects of the guideline's methodological quality were acceptable. At the recommendation level, the main implementation obstacles were the lack of explicit statements on the quality of the underlying evidence, lack of clarity around the strength of recommendations, and suboptimal computability hampering operationalisation of recommendations in electronic systems.

The scores for the NHG AF guideline were high in comparison with those assigned in other, similar guideline appraisal studies. For example, the systematic review of guideline appraisal studies by Alonso-Coello et al.[23] summarised the methodological quality of 626 guidelines, and reported mean AGREE II scores that were lower for all six domains. The study by Sabharwal et al.[17] appraised 101 cardiac clinical practice guidelines. Compared with the NHG AF guideline, they found higher mean scores for the 'Scope and Purpose' domain (64% and 85%, respectively) and for 'Clarity of Presentation' (67% versus 82%), but lower scores for the four remaining domains.

Suggestions to improve the NHG AF guideline

1. Extend stakeholder involvement in the guideline development process

The AGREE II domain 'Stakeholder involvement' obtained the lowest score of all domains (53%). This was partly due to the lack of representation of a wide range of stakeholders in the guideline development group, which consisted of general practitioners and an epidemiologist. Other relevant disciplines, such as neurologists and representatives of anticoagulation clinics, were consulted at the external review stage. Yet, AGREE II and other accepted guideline standards advocate multidisciplinary development groups because they tend to generate more balanced views than single-speciality groups.[24,25] Inviting representatives of other disciplines as members of the NHG guideline development group would be one way to increase multidisciplinary involvement. An alternative approach may be to consult the current panel of external reviewers earlier on in the process, for example when setting the guideline scope, selecting or rating the evidence, or when formulating the recommendations.

The low 'Stakeholder involvement' score also stemmed from the apparent limited extent to which patient experiences and expectations informed the NHG AF guideline. The national patient organisation was invited for external review, but it was unclear if and how their suggestions were addressed. Similar to the approaches for increased stakeholder involvement described above, patient involvement may be facilitated by having patient representation in the guideline development group, or by formal patient consultation at earlier stages of guideline development.[26] Alternatively, a literature review or patient interviews could inform a guideline section summarising patient views on and experiences with antithrombotic treatment for stroke prevention.

2. Standardise layout and language of key recommendations

Although the key recommendations are clearly presented in the online summary of the guideline (https://www.nhg.org/standaarden/samenvatting/atriumfibrilleren), they remain hidden in the full text version. Using a specific font, framing them in a box, or (if possible) presenting them as a flowchart would support distinguishing key recommendations from other types of information in the guideline. Additionally, applying standardised language may help guideline users to recognise recommendations as such, whilst aligning interpretations of whether recommendations should be considered relevant for all patients. The Grading of Recommendations, Assessment, Development and Evaluation (GRADE) framework[27] proposes standardised guideline language: recommendations can be either strong (level '1') or weak (level '2'), which translate into the phrases 'we recommend' or 'we suggest', respectively. Adopting GRADE as part of the guideline development methodology also ensures systematic assessment of the quality of the underlying evidence, with the letters 'A' (indicating high quality) to

'D' (very low quality) explicitly communicating evidence quality at the recommendation level. For future versions of the AF guideline, the NHG's updated procedure booklet (published January 2015) includes guidance on how to use the GRADE framework for assessing and summarising quality of the underlying evidence.

3. Suggest criteria for monitoring the guideline's use in practice
Facilitating local or regional monitoring of the guideline's use in practice requires clearly defined criteria derived from the guideline's key recommendations. For the NHG AF guideline, examples of criteria may be 'the percentage of female patients who are prescribed oral anticoagulants and are aged 65 years or older', or 'the percentage of patients on coumarin derivatives with an INR between 2.0 and 3.0'. Suggesting monitoring criteria as part of the guideline would provide a suitable starting point for developing audit and feedback, which Dutch GPs considered an encouraging strategy to improve guideline adherence.[28]

4. Prepare electronic implementation in parallel with the guideline development process
Recent studies have focused on providing GPs with clinical decision support systems (CDSS) to improve primary care stroke prevention in AF patients.[29-31] These systems use decision rules to evaluate the current treatment of the patient and, if necessary, recommend the GP to modify it. Creating these decision rules involves translating guidelines into a format that is interpretable by a computer. The GLIA dimensions 'Decidability' and 'Computability' relate to obstacles for electronic implementation, i.e. translating guideline recommendations into actionable, computable decision rules. In the current study, the NHG AF guideline scored poorly for these dimensions, indicating presence of ambiguous terms. Although this guideline was included in CDSSs for Dutch GPs,[26,29] the suboptimal computability hampered interpretation and translation of individual guideline statements into electronic decision rules. Especially the lack of clear definitions for certain contraindications and the unavailability of structured data to identify contraindications in electronic health records required input from an expert group of clinicians to fully convert the guideline into unambiguous decision rules.[26] Based on this finding we suggest involving a CDSS specialist when formulating recommendations for future updates of the NHG AF guideline. By preparing electronic implementation in parallel with the guideline development process, vague and inconsistent recommendations can be identified and resolved before publication.[30] This may not only improve overall implementability of the guideline in practice, but also facilitate the development of effective CDSS interventions.

In conclusion, this study provides pointers for optimising future versions of the NHG AF guideline. Future research should investigate whether applying these suggestions indeed positively affects implementation of the guideline in primary care, which in turn may improve the adequacy of antithrombotic treatment and stroke prevention in patients with atrial fibrillation.

ACKNOWLEDGMENTS

We would like to thank Maureen van der Donk and Wim Opstelten for providing additional information on the NHG guideline development process.

DISLOSURES

Sabine van der Veer is funded by the European Renal Association – European Dialysis Transplant Association (ERA-EDTA) for a research fellowship on guideline development and implementation. Wim Lucassen is appointed member of the Authorisation Committee of the NHG.

REFERENCES

1. Heeringa J, van der Kuip DA, Hofman A, et al. Prevalence, incidence and lifetime risk of atrial fibrillation: the Rotterdam study. Eur Heart J. 2006;27:949-53.
2. Hart RG, Pearce LA, Aguilar MI. Meta-analysis: Antithrombotic therapy to prevent stroke in patients who have nonvalvular atrial fibrillation. Ann Intern Med. 2007;146:857-67.
3. Camm AJ, Lip GYH, De Caterina R, et al. 2012 focused update of the ESC Guidelines for the management of atrial fibrillation. Eur Heart J. 2012;33:2719-47.
4. NHG werkgroep atriumfibrilleren. M79 NHG-Standaard Atriumfibrilleren. 2013;56:392-401.
5. Nieuwlaat R, Olsson SB, Lip GYH, et al. Guideline-adherent antithrombotic treatment is associated with improved outcomes compared with undertreatment in high-risk patients with atrial fibrillation. The Euro Heart Survey on Atrial Fibrillation. Am Heart J. 2007;153:1006-12.
6. Arts DL, Visscher S, Opstelten W, Korevaar JC, Abu-Hanna A, van Weert HC. Frequency and risk factors for under- and over-treatment in stroke prevention for patients with non-valvular atrial fibrillation in general practice. PLoS One. 2013;8:e67806.
7. Cabana MD, Rand CS, Powe NR, et al. Why don't physicians follow clinical practice guidelines? A framework for improvement. JAMA. 1999;282:1458-65.
8. Shekelle PG, Kravitz RL, Beart J, Marger M, Wang M, Lee M. Are nonspecific practice guidelines potentially harmful? A randomized comparison of the effect of nonspecific versus specific guidelines on physician decision making. Heal Serv Res. 2000;34:1429-48.
9. Lugtenberg M, Zegers-van Schaick JM, Westert GP, Burgers JS. Why don't physicians adhere to guideline recommendations in practice? An analysis of barriers among Dutch general practitioners. Implement Sci. 2009;4:54.
10. Woolf S, Schünemann HJ, Eccles MP, Grimshaw JM, Shekelle P. Developing clinical practice guidelines: types of evidence and outcomes; values and economics, synthesis, grading, and presentation and deriving recommendations. Implement Sci. 2012;7:61.

11. Gagliardi AR, Brouwers MC, Palda VA, Lemieux-Charles L, Grimshaw JM. How can we improve guideline use? A conceptual framework of implementability. Implement Sci. 2011;6:26.
12. NHG. Het NHG ter introductie [Internet]. 2016 [cited 2016 Mar 16]. Available from: www.nhg.org/het-nhg-ter-introductie
13. Van Hassel DTP, Kaseleijn A, Kenens RJ. Cijfers uit de registratie van huisartsen. Utrecht, Nivel: 2014.
14. Heiligers PJM, Noordman J, Korevaar JC, et al. Kennisvraag: praktijkondersteuners in de huisartspraktijk (POH's) klaar voor de toekomst? Utrecht: Nivel, 2012.
15. Brouwers MC, Kho ME, Browman GP, et al. AGREE II: advancing guideline development, reporting and evaluation in health care. Can Med Assoc J. 2010;182:E839-42.
16. Shiffman RN, Dixon J, Brandt C, et al. The GuideLine Implementability Appraisal (GLIA): development of an instrument to identify obstacles to guideline implementation. BMC Med Inform Decis Mak. 2005;5:23.
17. Sabharwal S, Patel V, Nijjer SS, et al. Guidelines in cardiac clinical practice: evaluation of their methodological quality using the AGREE II instrument. J R Soc Med. 2013;106:315-22.
18. Van Dijk LJ, Nelen WL, D'Hooghe TM, et al. The European Society of Human Reproduction and Embryology guideline for the diagnosis and treatment of endometriosis: an electronic guideline implementability appraisal. Implement Sci. 2011;6:7.
19. Nagler EV, Vanmassenhove J, van der Veer SN, et al. Diagnosis and treatment of hyponatremia: a systematic review of clinical practice guidelines and consensus statements. BMC Med. 2014;12:1.
20. Brouwers MEA. Appraisal of Guidelines for Research & Evaluation II. Agree Next Steps Consort. 2009;(May):1-56.
21. Workgroup eGLIA. eGLIA 2 instructions [Internet]. 2011. Available from: nutmeg.med.yale.edu/egliа2/eGLIAInstructions.pdf
22. NHG. NHG Procedure booklet [Internet]. 2010 [cited 2015 Jan 20]. Available from: https://www.nhg.org/sites/default/files/content/nhg_org/uploads/procedureboek_versie_090810.pdf
23. Alonso-Coello P, Irfan A, Solà I, et al. The quality of clinical practice guidelines over the last two decades: a systematic review of guideline appraisal studies. Qual Saf Health Care. 2010;19:e58.
24. Graham R, Mancher M, Wolman DM, Greenfiled S, Steinberg E. Institute of Medicine: Clinical Practice Guidelines We Can Trust. Natl Acad Press. 2011;1-13.
25. Eccles MP, Grimshaw JM, Shekelle P, Schünemann HJ, Woolf S. Developing clinical practice guidelines: target audiences, identifying topics for guidelines, guideline group composition and functioning and conflicts of interest. Implement Sci. 2012;7:60.
26. (G-I-N) GIN. G-I-N PUBLIC Toolkit : Patient and public involvement in guidelines. [Internet]. 2012. Available from: http://www.g-i-n.net/working-groups/gin-public/toolkit
27. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: An emerging consensus on rating quality of evidence and strength of recommendations. Chinese J Evidence-Based Med. 2009;9:8-11.
28. Lugtenberg M, Burgers JS, Han D, Westert GP. General practitioners' preferences for interventions to improve guideline adherence. J Eval Clin Pract. 2014;20:820-6.
29. Arts DL, Abu-Hanna A, Büller HR, Peters RJG, Eslami S, van Weert HCPM. Improving stroke prevention in patients with atrial fibrillation. Trials. 2013;14:193.
30. Chen R, Valladares C, Corbal I, Anani N, Koch S. Early Experiences from a guideline-based computerized clinical decision support for stroke prevention in atrial fibrillation. Stud Health Technol Inform. 2013;192:244-7.
31. Bajorek B, Magin P, Hilmer S, Krass I. A cluster-randomized controlled trial of a computerized antithrombotic risk assessment tool to optimize stroke prevention in general practice: a study protocol. BMC Health Serv Res. 2014;14:55.
32. Goud R, Hasman A, Strijbis A-M, Peek N. A parallel guideline development and formalization strategy to improve the quality of clinical practice guidelines. Int J Med Inform. 2009;78:513-20.

Appendix 1. *Individual appraisers' scores for AGREE II appraisal after group discussion; scores prior to group discussion are given in round brackets if they differed from scores after the discussion*

Appraiser					**Obtained domain score**
Domain and items	**1**	**2**	**3**	**4**	
SCOPE AND PURPOSE					
Overall objective of the guideline is specifically described	6	6	3	4	58
Health questions covered by the guideline are specifically described	4	3	3	6	
Target patient population of the guideline is specifically described	5	5	6	7	
STAKEHOLDER INVOLVEMENT					
All relevant professional groups were included in the guideline development group	3	5	4 (6)	4 (6)	50
Views and preferences of the target patient population were sought	2	3 (5)	3 (5)	3 (5)	
Target users of the guideline are clearly defined	6	7	5	5	
RIGOUR OF DEVELOPMENT					
Systematic methods were used to search for evidence	5 (2)	6	4 (1)	4	153
Criteria for selecting the evidence are clearly described	1	3	4 (5)	2	
Strengths and limitations of the body of evidence are clearly described	3	5 (7)	3	1	
Methods for formulating recommendations are clearly described	1	3	3 (5)	3 (6)	
Health benefits, side effects, and risks have been considered	7	7	6	6 (5)	
There is an explicit link between recommendations and supporting evidence	5	5	5	7	
The guideline was externally reviewed by experts prior to publication	6	7	7	6	
Procedure for updating the guideline is provided	7 (2)	7	7 (1)	7 (3)	
CLARITY OF PRESENTATION					
Recommendations are specific and unambiguous	5	5	5	7	60
Different disease management options are clearly presented	6	5	6	7	
Key recommendations are easily identifiable	3	4	4 (7)	3 (7)	
APPLICABILITY					
Facilitators and barriers to application of the guideline are described	5	4 (2)	4 (3)	3	71
Guideline provides advice and tools for applying recommendations in practice	4	6	3	6	
Potential resource implications of applying recommendations have been considered	6	6	6 (1)	6	
The guideline presents monitoring and/or auditing criteria	3	3	2 (1)	4 (6)	
EDITORIAL INDEPENDENCE					
Views of funding body have not influenced the guideline content	6 (1)	7	6	5	50
Competing interests of authors have been recorded and addressed	6 (4)	7	7	6	

PHOTO QUIZ

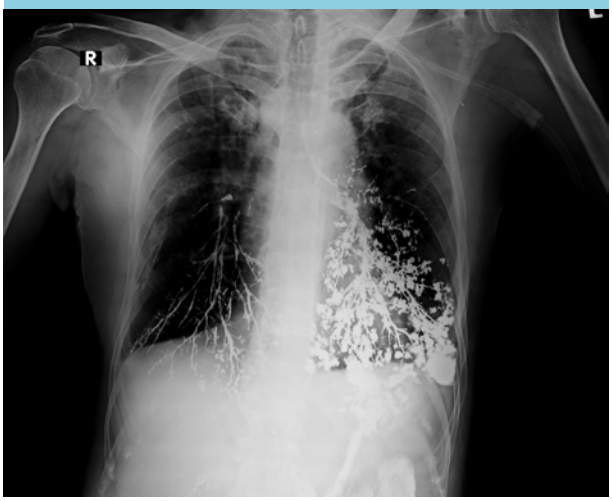
Acute-onset breathlessness after a radiological procedure

Y.K. Shejul*, R. Pendse, A. Kulkarni

Department of Medicine, BARC Hospital, Mumbai, India,
*corresponding author: tel.: +91 98 204 521 05, fax: 022 - 25506944, email: yogeshshejul@hotmail.com

A 76-year-old man with a background of chronic obstructive pulmonary disease presented with new-onset dysphagia. On examination, pooling of saliva in the pyriform fossa was noted. Systemic examination was otherwise unremarkable. He went on to have a barium swallow examination. During the procedure, he became breathless and oxygen saturation dropped to 85% on room air. Bilateral crackles and wheezing were noted on auscultation. He was transferred directly to our intensive care unit. Chest X-ray was performed immediately, which showed bilateral dense punctate foci and linear opacities in the alveoli and tracheobronchial tree, with greater involvement of the left side as compared with the right (*figure 1*). What is your diagnosis?

Figure 1. *Chest X-ray showing bilateral dense punctate foci and linear opacities, with greater involvement of the left side as compared to the right*



WHAT IS YOUR DIAGNOSIS?

See page 172 for the answer to this photo quiz.

ANSWER TO PHOTO QUIZ (PAGE 171)

ACUTE-ONSET BREATHLESSNESS AFTER A RADIOLOGICAL PROCEDURE

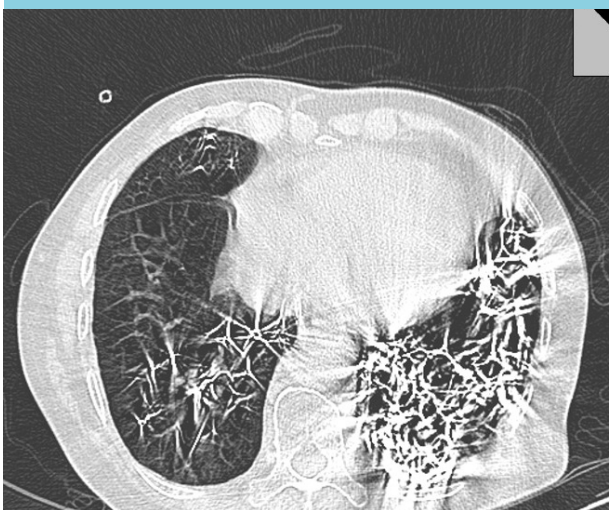
DIAGNOSIS

Chest CT revealed extensive areas of linear hyperdensities causing streaking artefacts along the tracheobronchial tree up to the level of the distal bronchioles (*figure 2*).

A diagnosis of barium aspiration was made on the basis of the history and radiological findings. The patient was managed with intravenous antibiotics (piperacillin/tazobactam and metronidazole), steroids (hydrocortisone), inhaled bronchodilators and supplemental oxygen, and received chest physiotherapy. He made a complete recovery and is on regular outpatient follow-up with no long-term sequelae at six months.

Barium swallow is a routine procedure in the evaluation of the oropharynx and oesophagus. The contrast medium commonly used is barium sulphate, an insoluble salt of barium. Aspiration of barium sulphate is a well-known complication. Risk factors include extremes of age, disordered swallowing, neuromuscular dysfunction, broncho-oesophageal fistula, alcoholism, head and neck cancer and functional gastrointestinal disorders. In our patient, the only risk factor for aspiration was his age. The extent and region of lung involvement depend on the position of the patient during and after aspiration. Diagnosis is usually made by chest X-ray or detection of barium during bronchoscopy. Barium is considered to be an inert material; however, acute respiratory distress, pneumonitis, sepsis and death have been reported after barium aspiration, especially with large volume aspirations.[1,2] The prognosis depends on the density and quantity of aspirated contrast, distribution of barium in the airways, the general condition of the patient and whether concomitant aspiration of gastric contents has occurred. There are no defined protocols for the management of barium aspiration. Bronchoscopic clearance of aspirated contrast is recommended in large volume aspiration and is aimed at removing as much barium as possible from the respiratory tract. Antibiotic therapy should be administered if associated infection is suspected and chest physiotherapy with postural drainage can assist in clearance of barium from the respiratory tract.[3] Iso-osmolar or low-osmolar contrast media can be used as a safe alternative to barium sulphate in patients at risk of aspiration, since they are associated with minimal morbidity or mortality risk if aspirated.[4]

Figure 2. *Chest CT showing areas of linear hyperdensities causing streaking artifacts*



REFERENCES

1. Fruchter O, Dragu R. A deadly examination. N Engl J Med. 2003;348:1016.
2. Gray C, Sivaloganathan S, Simpkins K. Aspiration of high-density barium contrast medium causing acute pulmonary inflammation–report of two fatal cases in elderly women with disordered swallowing. Clin Radiol. 1989;40:397-400.
3. Katsanoulas C, Passakiotou M, Mouloudi E, Gritsi-Gerogianni N, Georgopoulou V. Severe barium sulphate aspiration: a report of two cases and review of the literature. Signa Vitae. 2007;2:25-8.
4. Gelfand DW. Complications of gastrointestinal radiologic procedures: I. Complications of routine fluoroscopic studies. Gastrointest Radiol. 1980;5:293-315.

PHOTO QUIZ

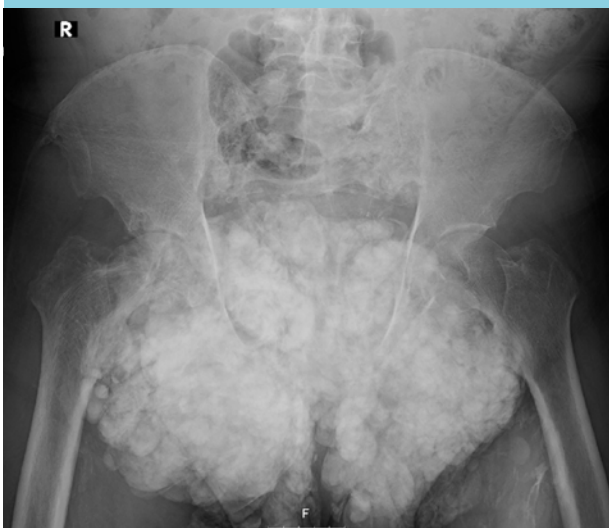
A haemodialysis patient with progressive leg pain

J. Hanssen[1]*, K. Berend[1], J. Tai[2], N. Vinke[3]

Departments of [1]Internal Medicine, [2]Radiology, [3]Orthopedics, St Elisabeth Hospital, Willemstad, Curaçao, *corresponding author: tel.: +599 946 249 00, fax: +599 946 247 39, email: jljhanssen@gmail.com

A 71-year-old man experienced progressive pain in his left upper leg and groin over a period of several months. He had been receiving haemodialysis for ten years because of end-stage renal disease due to chronic pyelonephritis. Medication included alfacalcidol and cinacalcet. On examination, multiple large tender rubber-like nodules were palpated in the symphysis and upper leg region. His monthly lab results usually showed slightly raised serum calcium and phosphate levels around 2.80 mmol/l and 1.90 mmol/l respectively. The parathyroid hormone level was generally above 1200 pg/ml (normal range 11-62 pg/ml) due to noncompliance. Pelvis radiography and computed tomography showed massive cauliflower-like calcifications with fluid levels in the pubic region, lesser pelvis and the proximal femora *(figure 1 and 2)*.

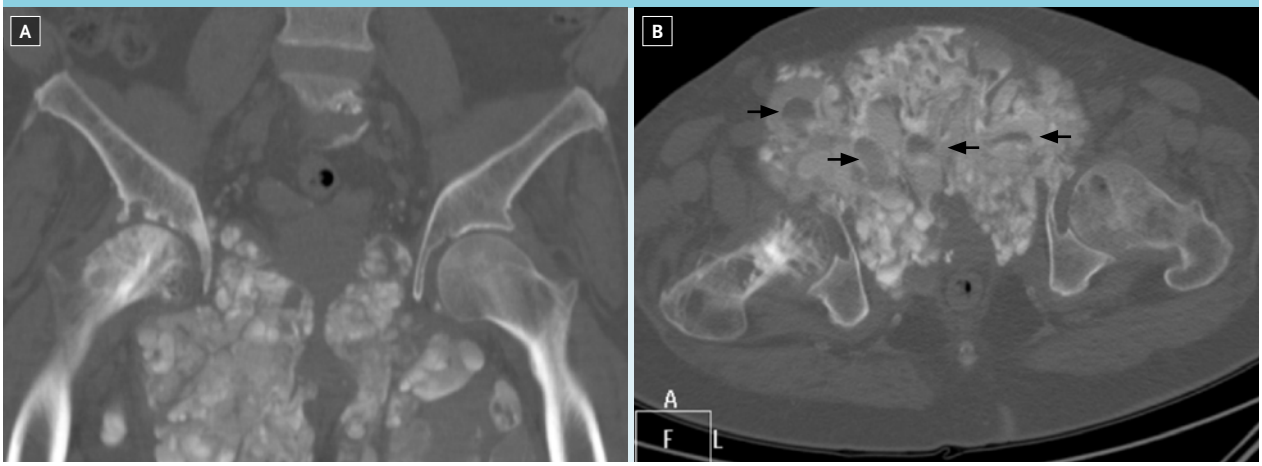
Figure 1. *Pelvis X-ray showing cauliflower-like calcifications*



WHAT IS YOUR DIAGNOSIS?

See page 174 for the answer to this photo quiz.

Figure 2. *CT pelvis. A. Sagital slide with calcificated cystic multilobulated lesions; B. Transversal slide with the same lesions also showing fluid levels (arrows)*



ANSWER TO PHOTO QUIZ (PAGE 173)

A HAEMODIALYSIS PATIENT WITH PROGRESSIVE LEG PAIN

DIAGNOSIS

The history of chronic kidney failure and secondary hyperparathyroidism together with these radiological findings are typical for secondary tumoral calcinosis. A biopsy, performed to exclude a malignancy, revealed non-specific calcified material without malignant-appearing cells. Tumoral calcinosis is a rare benign condition characterised by lobular calcified soft tissue deposits that can manifest around any joint. The hip, elbow and shoulder are the most common locations.[1] It is associated with hyperphosphataemia but the exact aetiology is still being unravelled. Tumoral calcinosis may be classified as a) primary normo-phosphataemic tumoral calcinosis; b) primary hyperphosphataemic tumoral calcinosis, where gene mutations lead to reduced urinary phosphate excretion causing hyperphosphataemia; c) secondary tumoral calcinosis mostly seen in chronic renal failure associated with secondary or tertiary hyperparathyreodism.[2,3]

The diagnosis of tumoral calcinosis in patients with chronic renal failure is mainly made by radiological evaluation.[3] Imaging studies show distinct amorphous, multilobulated and cystic calcifications with fluid-fluid levels ('sedimentation sign') in periarticular locations.[2] Serum biochemical tests (calcium, phosphorus, renal function tests, parathormone level and 1,25-dihydroxy-vitamin D) can help to distinguish between the different forms of tumoral calcinosis and from diseases that can mimic it. A characteristic presentation does not usually require pathological evaluation. When malignancy is suspected, however, a biopsy is needed. Conditions that can mimic tumoral calcinosis are osteochondromatosis, synovial sarcoma and osteosarcoma.[3]

Treatment of tumoral calcinosis in renal failure is often challenging and includes calcium and phosphorus restricted diets, intensive dialysis treatment, phosphate binders and adequate management of secondary or tertiary hyperparathyroidism, tailored to individual cases. When these medical interventions remain unsuccessful, parathyroidectomy may be necessary. Kidney transplantation may be another treatment option, after correction of the hyperparathyroidism. Surgical treatment of the calcifications is generally not recommended in the secondary form while this remains the first choice of treatment for the two primary forms of tumoral calcinosis.[3] The patient was treated with a phosphate and calcium restricted diet and phosphate binders but was incompliant and he rejected parathyroidectomy. The condition of the patient deteriorated and he passed away. Autopsy was refused.

REFERENCES

1. Olsen KM, Chew FS. Tumoral calcinosis: pearls, polemics, and alternative possibilities. Radiographics. 2006;26:871-85.
2. Smack D, Norton SA, Fitzpatrick JE. Proposal for a pathogenesis-based classification of tumoral calcinosis. Int J Dermatol. 1996;35:65-71.
3. Fathi I, Sakr M. Review of tumoral calcinosis: A rare clinico-pathological entity. World J Clin Cases. 2014;2:409-14.

PHOTO QUIZ

Things are not always what they seem (other causes of hepato-splenic nodules)

R. Caballero[1]*, D. Ibañez[2], N. Yanguas[2], M.F. Pérez[3], G. Aísa[4]

Departments of [1]Internal Medicine, [2]Radiology, [3]General Surgery, Hospital Reina Sofía, Tudela (Navarra), Spain, [4]Service Laboratory, Microbiology and Pathology, Hospital Reina Sofía, Tudela (Navarra), Spain, *corresponding author: tel.: +34 - 649 888 712, email: rcaballeroasensio@gmail.com

CASE REPORT

We present the case of a 43-year-old obese patient with type 2 diabetes. The abdominal ultrasonography performed in the context of an acute episode of renal colic showed multiple hypoechogenic hepatic lesions and an extended examination was required. A thoracic-abdominal CT scan was performed (*figure 1*) confirming the existence of multiple hepatic and splenic injuries and small thoracic and retroperitoneal lymphadenopathies.

Blood tests were performed and we found high levels of beta-2 microglobulin (> 3000 units), with all other results falling within the normal limits. The serology, autoantibodies, Mantoux and QuantiFERON tests were negative. A positron emission tomography was performed which showed lesions in the liver and spleen and multiple lymph nodes that had significantly increased fluordeoxyglucose uptake as compared to the surrounding parenchyma.

In order to reach a diagnosis, a splenectomy was scheduled. The result of the splenectomy (*figure 2*) was a piece of spleen of almost 900 g and 18 x 11 x 6.5 cm with multiple white superficial nodules, indurated, which represented 75% of the splenic parenchyma. On biopsy multiple granulomas of different sizes consisting of epithelioid cells, multinucleated Langerhans cells and small lymphocytes could be observed. It did not show necrosis. They were accompanied by vascular chronic inflammatory infiltrate. These findings were also present in ten lymph nodes in the splenic hilum. In addition, high levels of angiotensin-converting enzyme (132 U/l, normal range 13-64) were detected.

WHAT IS YOUR DIAGNOSIS?

See page 176 for the answer to this photo quiz.

Figure 1. *Multiple focal hypointense splenic lesions on T1 in the arterial phase of an MRI axial examination*

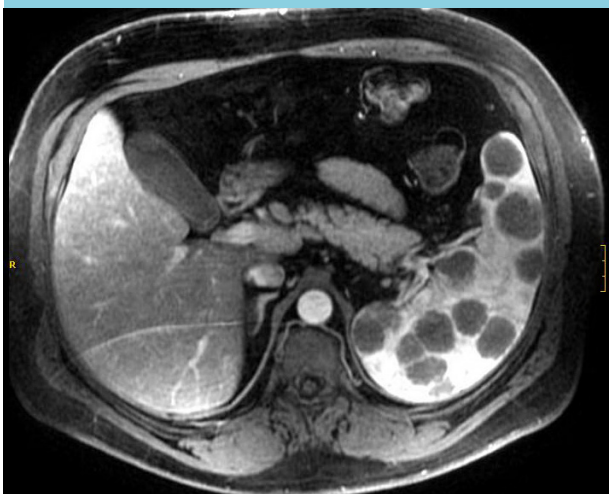
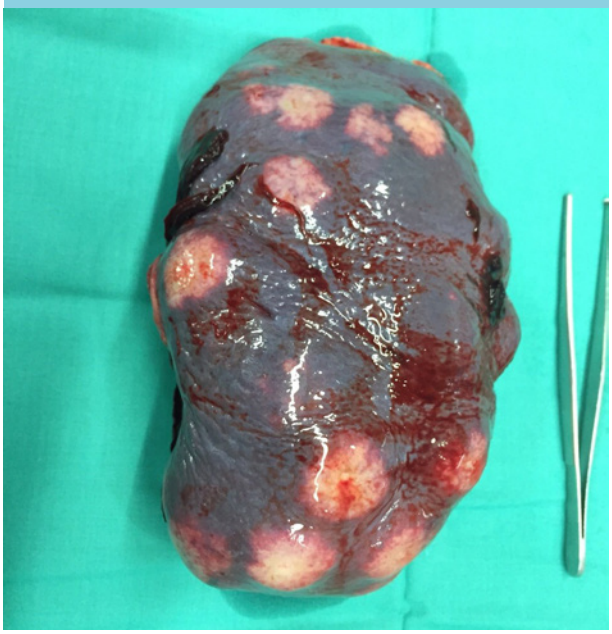


Figure 2. *Part of splenectomy with multiple lesions similar to those shown on CT and MRI*



ANSWER TO PHOTO QUIZ (PAGE 175)

THINGS ARE NOT ALWAYS WHAT THEY SEEM (OTHER CAUSES OF HEPATO-SPLENIC NODULES)

DIAGNOSIS

The final clinical diagnosis was hepato-splenic sarcoidosis (extrapulmonary disease). Because of the elevated liver transaminases, oral corticosteroid treatment was initiated at a dose of 0.5-1 mg/kg/day, with good tolerance and clinical evolution in the patient.

Sarcoidosis is a multisystem granulomatous disease of unknown aetiology, where diagnosed patients have pulmonary involvement in almost 90% of the cases.[1] Extrapulmonary involvement is common and all organs may be affected (especially the lymph nodes, eyes, joints, central nervous system), but it is rare to find isolated extrapulmonary disease (less than 10% of the patients).[2] Without any symptoms, it is even more difficult to make the diagnosis. The isolated splenic or hepatosplenic involvement is uncommon, and these patients may not have any symptoms at all. In this situation it may be necessary to use multiple tests because there may be added difficulty in differentiating the lesions of other diseases such as lymphoma or metastasis[3], or even reach splenectomy.[4]

REFERENCES

1. Lannuzzi MC, Rybicki BA, Teirstein AS. Sarcoidosis. N Engl J Med. 2007;357: 2153-65.
2. Vardhanabhuti V, Venkatanarasimha N, Bhatnagar G, et al. Extra-pulmonary manifestations of sarcoidosis. Clin Radiol. 2012;67:263-76.
3. Agrawal K, Chawla YK, Bhattacharya A, Mittal BR. Fluorodeoxyglucose positron emission tomography / computed tomography findings in nodular hepatic and splenic sarcoidosis. World J Nucl Med. 2014;13:144-5.
4. Raber EL, Bean J, Beck P. Splenic sarcoidosis: a case report and review of the imaging findings of multiple incidental splenic lesions as the initial presentation of sarcoidosis. Can J Gastroenterol. 2011;25:477-8.

LETTER TO THE EDITOR

Leptospiral meningitis in adults

V.M. dos Santos

Department of Internal Medicine, Armed Forces Hospital and Catholic University of Brasília, Brasília-DF, Brazil, tel.: +55-61 32330812, fax: +55-61 32331599, email: vitorinomodesto@gmail.com

Dear Editor,

Although leptospirosis is not a major public health problem in the Netherlands, its incidence as well the rate of infections locally contracted (62%) have increased more recently, as I read in the interesting article by Van Samkar and colleagues about leptospiral meningitis.[1] They reviewed 19 adults, 78.8% males, in the age range of 17-61 years, and 95% of the individuals contracted leptospirosis outside the Netherlands. Suspicion of meningitis occurred in seven cases and was confirmed in four patients (21%).[1] Moreover, the authors reviewed 366 cases of leptospiral meningitis affecting adult people from Europe, Asia and America, published between 1947 and 2014. Similarities were observed in male prevalence (82%) and age range (17-77 years).[1] Major clinical features and laboratory data related to diagnoses of leptospirosis and leptospiral meningitis, and the favourable role of early antibiotic therapy on the outcome of severe complications were emphasised.[1-4] Contact with contaminated fresh or sewage water was highlighted as a risk factor for leptospiral infection.[1-4] While the estimated incidence of leptospirosis in the Netherlands is 0.57 per 100,000 inhabitants,[1] the globally estimated incidence is approximately 500,000 severe cases yearly.[4] Weil's syndrome includes jaundice, renal failure and haemorrhagic phenomena, and represents the most severe clinical manifestations of leptospirosis, involving high mortality.[2-4] Weil's syndrome was not observed in the four patients with meningitis described by Van Samkar and colleagues.

In this setting, I would like to comment on the case report of a 19-year-old Brazilian man with the typical syndrome in addition to leptospiral meningitis successfully controlled by penicillin.[3] The infection caused by *L. grippotyphosa* was associated with swimming in a river barrage. Laboratory determinations showed elevated blood levels of urea nitrogen, creatinine, and total bilirubin, and the routine tests on the cerebrospinal fluid confirmed the diagnosis of aseptic meningitis. This young man had severe disturbance of hepatic and renal functions, in addition to conspicuous cutaneous and conjunctival haemorrhagic phenomena. The improvement was due to early diagnostic suspicion and administration of intravenous penicillin G. The patient's hospital discharge occurred three weeks after admission, without any sequelae.[3] Although icteric patients are more prone to have poorer outcomes, anicteric leptospirosis has also been associated with complications such as pneumonitis, pancreatitis and pericarditis.[2,4]

DISCLOSURES

The author declares no conflict of interest. No funding or financial support was received.

REFERENCES

1. Van Samkar A, van de Beek D, Stijnis C, Goris M, Brouwer MC. Suspected leptospiral meningitis in adults: report of four cases and review of the literature. Neth J Med. 2015;73:464-70.
2. Helmerhorst HJ, van Tol EN, Tuinman PR, et al. Severe pulmonary manifestation of leptospirosis. Neth J Med. 2012;70:215-21.
3. Dos Santos VM, dos Santos JA, Sugai TA, dos Santos LA. Weil's syndrome. Rev Cubana Med Trop. 2003;55:44-6.
4. Santos VM, Santos UM, Gebrin DG, Santos AM, Cancado AC. Anicteric leptospirosis with pneumonitis, pericarditis and acalculous cholecystitis. Infez Med. 2014;22:236-40.

LETTER TO THE EDITOR

Behçet's disease: ethnicity and associated conditions

V.M. dos Santos

Department of Internal Medicine, Armed Forces Hospital and Catholic University of Brasília, Brasília-DF, Brazil, tel.: +55-61 32330812, fax: +55-61 32331599, email: vitorinomodesto@gmail.com

Dear Editor,

Behçet's disease presents with recurrent oral ulceration, genital ulcers, ocular inflammation and skin lesions, and arthritis, neurological and gastrointestinal inflammations.[1-3] I read with interest the article by Kappen and colleagues about the prevalence and manifestations of Behçet's disease, and the correlation of severity and morbidity with patient ethnicity.[1] They studied 84 patients with Behçet's disease of Dutch, Turkish, and Moroccan descent. The prevalence was similar to the countries of ancestry: 1, 39 and 71 per 100,000 for Dutch, Moroccan and Turkish descendants, respectively; and the same was observed about morbidity.[1] Notwithstanding, both uveitis and pustules occurred with more frequency in the Netherlands.[1]

I would like to highlight findings of the Brazilian review by Oliveira and colleagues including 60 Behçet's disease patients with a mean age of 40 (SD, 10.7) years, and with female-male ratio 1.2:1.0. The total population of Dutch, Turkish and Moroccan descendants is not very large in Brazil; moreover, all patients were born in this country and none were related to specific ethnic groups. As the distribution by ethnicity only revealed 55% mestizos, 36.7% whites, and 8.3% blacks,[2] the aim is to compare clinical features in Brazilian patients with those cited by Kappen et al.[1] The frequency of clinical manifestations was oral ulcers (100%), genital ulcers (93.3%), cutaneous lesions (71.7%), ocular manifestations (63.3%), arthritis (46.7%), neurological involvement (28.3%), pathergy test (22.7%), thrombosis (13.3%) and gastrointestinal involvement (3.3%).[2] Worthy of note, arthritis and erythema nodosum occurred more often in women and papulopustular lesions in men, and prognosis of the entire group was favourable.[2] Behçet's disease has been described in association with heart changes, thromboembolism, and colon tumours, and the concomitance of this condition with patent *foramen ovale*, antiphospholipid antibodies syndrome, and colon polyps was described in a 49-year-old Brazilian woman.[3] Thromboembolic phenomena did not occur in this lady, but cerebral thromboembolism was described in a Japanese man with Behçet's disease and a similar cardiac anomaly.[4] Cases of Behçet's disease are infrequent in Brazil as in the Netherlands, and few studies have evaluated significant numbers of cases with a confirmed diagnosis.[2,3] Therefore, less severe cases in patients with only recurrent oral ulcerations may be underdiagnosed by primary care physicians. The author believes that the articles discussed here may enhance the suspicion index about Behçet's disease and associated disorders, contributing to reducing misdiagnosis and underdiagnosis, mainly in regions where the ethnic groups under high risk of Behçet's disease are not well characterised.

DISCLOSURES

The author declares no conflict of interest. No funding or financial support was received.

REFERENCES

1. Kappen JH, van Dijk EHC, Baak-Dijkstra M, et al. Behçet's disease, hospital-based prevalence and manifestations in the Rotterdam area. Neth J Med. 2015;73:471-7.
2. Oliveira ACD, Buosi ALP, Dutra LA, de Souza AWS. Behçet disease: clinical features and management in a Brazilian tertiary hospital. J Clin Rheumatol. 2011;17:416-20.
3. Santos VM, Carvalho MRM, Silva FHO, Cançado ACV, Guimarães JPF, Zilcem da Costa Arruda Junior ZC. [Behçet's disease, anticardiolipin antibodies, patent foramen ovale, and Sweet syndrome – an uncommon association]. Brasília Med. 2013;50:261-5.
4. Ishida C, Furui E, Sakashita Y, Yamada M. Embolic stroke with a patent foramen ovale and Behçet disease. Intern Med. 2005;44:326-7.